Quantum Mechanics

Richard Fitzpatrick

Professor of Physics
The University of Texas at Austin

Major Sources

The textbooks that I have consulted most frequently while developing course material are:

The Principles of Quantum Mechanics. P.A.M. Dirac, 4th Edition (revised) (Oxford University Press, Oxford UK, 1958).

***The Feynman Lectures on Physics*.** R.P. Feynman, R.B. Leighton, and M. Sands, Volumes I–III (Addison-Wesley, Reading MA, 1965).

***Quantum Mechanics*.** E. Merzbacher, 2nd Edition (John Wiley & Sons, New York NY, 1970).

***Quantum Physics*.** S. Gasiorowicz, 2nd Edition (John Wiley & Sons, New York NY, 1996).

***Modern Quantum Mechanics*.** J.J. Sakurai, and J. Napolitano, 2nd Edition (Addison-Wesley, Boston MA, 2011).

Purpose of Course

Quantum mechanics was developed during the first few decades of the twentieth century via a series of inspired guesses made by various physicists, including Planck, Einstein, Bohr, Schrödinger, Heisenberg, Pauli, and Dirac. All of these scientists were trying to construct a self-consistent theory of microscopic dynamics that was compatible with experimental observations. The purpose of this course is to present quantum mechanics in a systematic fashion, starting from the fundamental postulates, and developing the theory in as logical a manner as possible.

1 Fundamental Concepts

1.1 Breakdown of Classical Physics

The necessity for a departure from classical physics is demonstrated by the following phenomena:

1. **Anomalous Atomic and Molecular Stability.** According to classical physics, an electron orbiting an atomic nucleus undergoes acceleration and should, therefore, lose energy via the continuous emission of electromagnetic radiation, causing it to gradually spiral in towards the nucleus. Experimentally, this is not observed to happen.

2. **Anomalously Low Atomic and Molecular Specific Heats.** According to the equipartition theorem of classical physics, each degree of freedom of an atomic or molecular system should contribute $R/2$ to its molar specific heat capacity, where R is the molar ideal gas constant. In fact, only the translational, and some rotational, degrees of freedom seem to contribute. The vibrational degrees of freedom appear to make no contribution at all (except at high temperatures). Incidentally, this fundamental problem with classical physics was known and appreciated by the middle of the nineteenth century. Stories that physicists at the start of the twentieth century thought that classical physics explained everything, and that there was nothing left to discover, are largely apocryphal (see Feynman, Volume I, Chapter 40).

3. **Ultraviolet Catastrophe.** According to classical physics, the equilibrium energy density of an electromagnetic field contained within a vacuum cavity whose walls are held at a fixed temperature is infinite, due to a divergence of energy carried by short wavelength modes. This divergence is called the *ultraviolet catastrophe*. Experimentally, there is no such divergence, and the total energy density is finite.

4. **Wave-Particle Duality.** Classical physics treats waves and particles as completely distinct phenomena. However, various experiments (e.g., the interference of light, the photoelectric effect, electron diffraction) demonstrate that waves sometimes act as if they were streams of particles, and streams of particles sometimes act as if they were waves. This behavior is completely inexplicable within the framework of classical physics.

1.2 Photon Polarization

It is known experimentally that if plane polarized light is used to eject photo-electrons then there is a preferred direction of emission of the electrons. Clearly, the polarization properties of light, which are more usually associated with its wave-like behavior, also extend to its particle-like behavior. In particular, a polarization can be ascribed to each individual photon in a beam of light.

Consider the following well-known experiment. A beam of plane polarized light is passed through a polarizing film, which is normal to the beam's direction of propagation, and which has the property that it is only transparent to light whose plane of polarization lies perpendicular to its optic axis (which is assumed to lie in the plane of the film). Classical electromagnetic wave theory tells us that if the beam is polarized perpendicular to the optic axis then all of the light is transmitted, if the beam is polarized parallel to the optic axis then none of the light is transmitted, and if the light is polarized at an angle α to the axis then a fraction $\sin^2\alpha$ of the beam energy is transmitted. Let us try to account for these observations at the individual photon level.

A beam of light that is plane polarized in a certain direction is made up of a stream of photons which are each plane polarized in that direction. This picture leads to no difficulty if the plane of polarization lies parallel or perpendicular to the optic axis of the polarizing film. In the former case, none of the photons are transmitted, and, in the latter case, all of the photons are transmitted. But, what happens in the case of an obliquely polarized incident beam?

The previous question is not very precise. Let us reformulate it as a question relating to the result of some experiment that we could perform. Suppose that we were to fire a single photon at a polarizing film, and then look to see whether or not it emerges on the other side. The possible results of the experiment are that either a whole photon (whose energy is equal to the energy of the incident photon) is observed, or no photon is observed. Any photon that is transmitted though the film must be polarized perpendicular to the film's optic axis. Furthermore, it is impossible to imagine (in physics) finding part of a photon on the other side of the film. If we repeat the experiment a great number of times then, on average, a fraction $\sin^2\alpha$ of the photons are transmitted through the film, and a fraction $\cos^2\alpha$ are absorbed. Thus, given that the trials are statistically independent of one another, we must conclude that a photon has a *probability* $\sin^2\alpha$ of being transmitted as a photon polarized in the plane perpendicular to the optic axis, and a probability $\cos^2\alpha$ of being absorbed. These values for the probabilities lead to the correct classical limit for a beam containing a large number of photons.

Note that we have only been able to preserve the individuality of photons, in all cases, by abandoning the determinacy of classical theory, and adopting a fundamentally *probabilistic* approach. We have no way of knowing whether an individual obliquely polarized photon is going to be absorbed by, or transmitted through, a polarizing film. We only know the probability of each event occurring. This is a fairly sweeping statement. Recall, however, that the state of a photon is fully specified once its energy, direction of propagation, and polarization are known. If we imagine performing experiments using monochromatic light, normally incident on a polarizing film, with a particular oblique polarization, then the state of each individual photon in the beam is completely specified, and nothing remains to uniquely determine whether the photon is transmitted or absorbed by the film.

The above discussion about the results of an experiment with a single obliquely polarized photon incident on a polarizing film answers all that can be legitimately asked about what happens to the photon when it reaches the film. Questions as to what determines whether the photon is transmitted or not, or how it changes its direction of polarization, are illegitimate, because they do not relate to the outcome of a possible experiment. Nevertheless, some further description is needed, in order to allow the results of this experiment to be correlated with the results of other experiments that can be performed using photons.

The further description provided by quantum mechanics is as follows. It is supposed that a photon polarized obliquely to the optic axis can be regarded as being partly in a state of polarization parallel to the axis, and partly in a state of polarization perpendicular to the axis. In other words, the oblique polarization state is some sort of *superposition* of two states of parallel and perpendicular polarization. Because there is nothing special about the orientation of the optic axis in our experiment, we deduce that any state of polarization can be regarded as a superposition of two mutually perpendicular states of polarization. When we cause a photon to encounter a polarizing film, we are subjecting it to an observation. In fact, we are observing whether it is polarized parallel or perpendicular to the film's optic axis. The effect of making this observation is to force the photon entirely into a state of parallel or perpendicular polarization. In other words, the photon has to jump suddenly from being partly in each of these two states to being entirely in one or the other of them. Which of the two states it will jump into cannot be predicted, but is governed by probability laws. If the photon jumps into a state of parallel polarization then it is absorbed. Otherwise, it is transmitted. Note that, in this example, the introduction of indeterminacy into the problem is clearly connected with the act of observation. In other words, the indeterminacy is related to the inevitable disturbance of the system associated with the act of observation.

1.3 Fundamental Principles of Quantum Mechanics

There is nothing special about the transmission and absorption of photons through a polarizing film. Exactly the same conclusions as those outlined above are obtained by studying other simple experiments, such as the interference of photons (see Dirac, Section I.3), and the Stern-Gerlach experiment (see Sakurai, Chapter 1; Feynman, Chapter 5). The study of these simple experiments leads us to formulate the following fundamental principles of quantum mechanics:

1. **Dirac's Razor.** Quantum mechanics can only answer questions regarding the outcome of possible experiments. Any other questions lie beyond the realms of physics.

2. **Principle of the Superposition of States.** Any microscopic system (i.e., an atom, molecule, or particle) in a given state can be regarded as being partly in each of two or more other states. In other words, any state can be regarded as a superposition of two or more other states. Such superpositions can be performed in an infinite number of different ways.

3. **Principle of Indeterminacy.** An observation made on a microscopic system causes it to jump into one or more particular states (which are related to the type of observation). It is impossible to predict into which final state a particular system will jump. However, the probability of a given system jumping into a given final state can be predicted.

The first of these principles was formulated by quantum physicists (such as Dirac) in the 1920's to fend off awkward questions such as "How can a system suddenly jump from one state into another?", or "How does a system decide which state to jump into?". As we shall see, the second principle is the basis for the mathematical formulation of quantum mechanics. The final principle is still rather vague. We need to extend it so that we can predict which possible states a system can jump into after a particular type of observation, as well as the probability of the system making a particular jump.

1.4 Ket Space

Consider a microscopic system composed of particles, or bodies, with specific properties (mass, moment of inertia, etc.) interacting according to specific laws of force. There will be various possible motions of the particles, or bodies, consistent with these laws of force. Let us term each such motion a *state* of the system. According to the principle of superposition of states, any given state can be regarded as a superposition of two or more other states. Thus, states must be related to mathematical quantities of a kind that can be added together to give other quantities of the same kind. The most obvious examples of such quantities are *vectors*.

Let us consider a particular microscopic system in a particular state, which we label A: e.g., a photon with a particular energy, momentum, and polarization. We can represent this state as a particular vector, which we also label A, residing in some vector space, where the other elements of the space represent all of the other possible states of the system. Such a space is called a *ket space* (after Dirac). The state vector A is conventionally written

$$|A\rangle. \tag{1.1}$$

Suppose that state A is, in fact, the superposition of two different states, B and C. This interrelation is represented in ket space by writing

$$|A\rangle = |B\rangle + |C\rangle, \tag{1.2}$$

where $|B\rangle$ is the vector relating to the state B, etc. For instance, state $|B\rangle$ might represent a photon propagating in the z-direction, and plane polarized in the x-direction, and state $|C\rangle$ might represent a similar photon plane polarized in the y-direction. In this case, the sum of these two states represents a photon whose plane of polarization makes an angle of 45° with both the x- and y-directions (by analogy with classical physics). This latter state is represented by $|B\rangle + |C\rangle$ in ket space.

Suppose that we want to construct a state whose plane of polarization makes an arbitrary angle α with the x-direction. We can do this via a suitably weighted superposition of states B and C. By analogy with classical physics, we require $\cos\alpha$ of state B, and $\sin\alpha$ of state C. This new state is represented by

$$\cos\alpha\,|B\rangle + \sin\alpha\,|C\rangle \tag{1.3}$$

in ket space. Note that we cannot form a new state by superposing a state with itself. For instance, a photon polarized in the y-direction superposed with another photon polarized in the y-direction (with the same energy and momentum) gives the same photon. This implies that the ket vector

$$c_1\,|A\rangle + c_2\,|A\rangle = (c_1 + c_2)\,|A\rangle \tag{1.4}$$

corresponds to the same state that $|A\rangle$ does. Thus, ket vectors differ from conventional vectors in that their magnitudes, or lengths, are physically irrelevant. All the states of the system are in one to one correspondence with all the possible directions of vectors in the ket space, no distinction being made between the directions of the ket vectors $|A\rangle$ and $-|A\rangle$. There is, however, one caveat to the above statements. If $c_1 + c_2 = 0$ then the superposition process yields nothing at all: i.e., no state. The absence of a state is represented by the null vector $|0\rangle$ in ket space. The null vector has the fairly obvious property that

$$|A\rangle + |0\rangle = |A\rangle \tag{1.5}$$

for any vector $|A\rangle$. The fact that ket vectors pointing in the same direction represent the same state relates ultimately to the quantization of matter: i.e., the fact that it comes in irreducible packets called photons, electrons, atoms, etc. If we observe a microscopic system then we either see a state (i.e., a photon, or an atom, or a molecule, etc.) or we see nothing—we can never see a fraction or a multiple of a state. In classical physics, if we observe a wave then the amplitude of the wave can take any value between zero and infinity. Thus, if we were to represent a classical wave by a vector then the magnitude, or length, of the vector would correspond to the amplitude of the wave, and the direction would correspond to the frequency and wavelength, so that two vectors of different lengths pointing in the same direction would represent different wave states.

We have seen, in Equation (1.3), that any plane polarized state of a photon can be represented as a linear superposition of two orthogonal polarization states in which the weights are real numbers. Suppose that we want to construct a circularly polarized photon state. Well, we know from classical physics that a circularly polarized wave is a superposition of two waves of equal amplitude, plane polarized in orthogonal directions, which oscillate in *phase quadrature*. This suggests that a circularly polarized photon is the superposition of a photon polarized in the x-direction (state B) and a photon polarized in the y-direction (state C), with equal weights given to the two states, but with the proviso that state C oscillates 90° out of phase with state B. By analogy with classical physics, we can use *complex numbers* to simultaneously represent the weighting and relative phase in a linear superposition. Thus, a circularly polarized photon is represented by

$$|B\rangle + \mathrm{i}\,|C\rangle \tag{1.6}$$

in ket space. A general elliptically polarized photon is represented by

$$c_1\,|B\rangle + c_2\,|C\rangle, \tag{1.7}$$

where c_1 and c_2 are complex numbers. We conclude that a ket space must be a complex vector space if it is to properly represent the mutual interrelations between the possible states of a microscopic system.

Suppose that the ket $|R\rangle$ is expressible linearly in terms of the kets $|A\rangle$ and $|B\rangle$, so that

$$|R\rangle = c_1\,|A\rangle + c_2\,|B\rangle. \tag{1.8}$$

We say that $|R\rangle$ is *dependent* on $|A\rangle$ and $|B\rangle$. It follows that the state R can be regarded as a linear superposition of the states A and B. So, we can also say that state R is dependent on states A and B. In fact, any ket vector (or state) that is expressible linearly in terms of certain others is said to be dependent on them. Likewise, a set of ket vectors (or states) are termed *independent* if none of them are expressible linearly in terms of the others.

The dimensionality of a conventional vector space is defined as the number of independent vectors contained in that space. Likewise, the dimensionality of a ket space is equivalent to the number of independent ket vectors it contains. Thus, the ket space that represents the possible polarization states of a photon propagating in the z-direction is two-dimensional (the two independent vectors correspond to photons plane polarized in the x- and y-directions, respectively). Some microscopic systems have a finite number of independent states (e.g., the spin states of an electron in a magnetic field). If there are N independent states then the possible states of the system are represented as an N-dimensional ket space. Some microscopic systems have a denumerably infinite number of independent states (e.g., a particle in an infinitely deep, one-dimensional, potential well). The possible states of such a system are represented as a ket space whose dimensions are denumerably infinite. Such a space can be treated in more or less the same manner as a finite-dimensional space. Unfortunately, some microscopic systems have a nondenumerably infinite number of independent states (e.g., a free particle). The possible states of such a system are represented as a ket space whose dimensions are nondenumerably infinite. This type of space requires a slightly different treatment to spaces of finite, or denumerably infinite, dimensions (see Section 1.15).

In conclusion, the states of a general microscopic system can be represented as a complex vector space of (possibly) infinite dimensions. Such a space is termed a *Hilbert space* by mathematicians.

1.5 Bra Space

A snack machine inputs coins plus some code entered on a key pad, and (hopefully) outputs a snack. It also does so in a deterministic manner: i.e., the same money plus the same code produces the same snack (or the same error message) time after time. Note that the input and output of the machine have completely different natures. We can imagine building a rather abstract snack machine which inputs ket vectors and outputs complex numbers in a deterministic fashion. Mathematicians call such a machine a *functional*. Imagine a general functional, labeled F, acting on a general ket vector, labeled A, and spitting out a general complex number ϕ_A. This process is represented mathematically by writing

$$\langle F|(|A\rangle) = \phi_A. \tag{1.9}$$

Let us narrow our focus to those functionals that preserve the linear dependencies of the ket vectors upon which they operate. Not surprisingly, such functionals are termed *linear functionals*. A general linear functional, labeled F, satisfies

$$\langle F|(|A\rangle + |B\rangle) = \langle F|(|A\rangle) + \langle F|(|B\rangle), \tag{1.10}$$

where $|A\rangle$ and $|B\rangle$ are any two kets in a given ket space.

Consider an N-dimensional ket space [i.e., a finite-dimensional, or denumerably infinite dimensional (i.e., $N \to \infty$), space]. Let the $|i\rangle$ (where i runs from 1 to N) represent N independent ket vectors in this space. A general ket vector can be written[1]

$$|A\rangle = \sum_{i=1,N} \alpha_i\,|i\rangle, \tag{1.11}$$

---

[1] Actually, this is only strictly true for finite-dimensional spaces. Only a special subset of denumerably infinite dimensional spaces have this property (i.e., they are complete). However, because a ket space must be complete if it is to represent the states of a microscopic system, we need only consider this special subset.

where the α_i are an arbitrary set of complex numbers. The only way that the functional F can satisfy Equation (1.10) for all vectors in the ket space is if

$$\langle F|(|A\rangle) = \sum_{i=1,N} f_i\,\alpha_i, \tag{1.12}$$

where the f_i are a set of complex numbers relating to the functional.

Let us define N basis functionals $\langle i|$ which satisfy

$$\langle i|(|j\rangle) = \delta_{i\,j}. \tag{1.13}$$

Here, the *Kronecker delta symbol* is defined such that $\delta_{i\,j} = 1$ if $i = j$, and $\delta_{i\,j} = 0$ otherwise. It follows from the previous three equations that

$$\langle F| = \sum_{i=1,N} f_i\,\langle i|. \tag{1.14}$$

But, this implies that the set of all possible linear functionals acting on an N-dimensional ket space is itself an N-dimensional vector space. This type of vector space is called a *bra space* (after Dirac), and its constituent vectors (which are actually functionals of the ket space) are called bra vectors. Note that bra vectors are quite different in nature to ket vectors (hence, these vectors are written in mirror image notation, $\langle\cdots|$ and $|\cdots\rangle$, so that they can never be confused). Bra space is an example of what mathematicians call a *dual vector space* (i.e., it is dual to the original ket space). There is a one to one correspondence between the elements of the ket space and those of the related bra space. So, for every element A of the ket space, there is a corresponding element, which it is also convenient to label A, in the bra space. That is,

$$|A\rangle \overset{\text{DC}}{\longleftrightarrow} \langle A|, \tag{1.15}$$

where DC stands for *dual correspondence*.

There are an infinite number of ways of setting up the correspondence between vectors in a ket space and those in the related bra space. However, only one of these has any physical significance. (See Section 1.11.) For a general ket vector A, specified by Equation (1.11), the corresponding bra vector is written

$$\langle A| = \sum_{i=1,N} \alpha_i^{\ast}\,\langle i|, \tag{1.16}$$

where the $\alpha_i^{\ast}$ are the complex conjugates of the α_i. $\langle A|$ is termed the dual vector to $|A\rangle$. It follows, from the above, that the dual to $c\,|A\rangle$ is $c^{\ast}\,\langle A|$, where c is a complex number. More generally,

$$c_1\,|A\rangle + c_2\,|B\rangle \overset{\text{DC}}{\longleftrightarrow} c_1^{\ast}\,\langle A| + c_2^{\ast}\,\langle B|. \tag{1.17}$$

Recall that a bra vector is a functional that acts on a general ket vector, and spits out a complex number. Consider the functional which is dual to the ket vector

$$|B\rangle = \sum_{i=1,N} \beta_i\,|i\rangle \tag{1.18}$$

acting on the ket vector $|A\rangle$. This operation is denoted $\langle B|(|A\rangle)$. Note, however, that we can omit the round brackets without causing any ambiguity, so the operation can also be written $\langle B||A\rangle$. This expression can be further simplified to give $\langle B|A\rangle$. According to Equations (1.11), (1.13), (1.16), and (1.18),

$$\langle B|A\rangle = \sum_{i=1,N} \beta_i^{\ast}\,\alpha_i. \tag{1.19}$$

Mathematicians term $\langle B|A\rangle$ the *inner product* of a bra and a ket.[2] An inner product is (almost) analogous to a scalar product between covariant and contravariant vectors in curvilinear coordinates. It is easily demonstrated that

$$\langle B|A\rangle = \langle A|B\rangle^{\ast}. \tag{1.20}$$

Consider the special case where $|B\rangle \to |A\rangle$. It follows from Equations (1.19) and (1.20) that $\langle A|A\rangle$ is a real number, and that

$$\langle A|A\rangle \geq 0. \tag{1.21}$$

The equality sign only holds if $|A\rangle$ is the null ket [i.e., if all of the α_i are zero in Equation (1.11)]. This property of bra and ket vectors is essential for the probabilistic interpretation of quantum mechanics, as will become apparent in Section 1.11.

Two kets $|A\rangle$ and $|B\rangle$ are said to be *orthogonal* if

$$\langle A|B\rangle = 0, \tag{1.22}$$

which also implies that $\langle B|A\rangle = 0$.

Given a ket $|A\rangle$, which is not the null ket, we can define a *normalized ket* $|\tilde{A}\rangle$, where

$$|\tilde{A}\rangle = \left(\frac{1}{\sqrt{\langle A|A\rangle}}\right)|A\rangle, \tag{1.23}$$

with the property

$$\langle \tilde{A}|\tilde{A}\rangle = 1. \tag{1.24}$$

Here, $\sqrt{\langle A|A\rangle}$ is known as the *norm* or “length” of $|A\rangle$, and is analogous to the length, or magnitude, of a conventional vector. Because $|A\rangle$ and $c\,|A\rangle$ represent the same physical state, it makes sense to require that all kets corresponding to physical states have unit norms.

It is possible to define a dual bra space for a ket space of nondenumerably infinite dimensions in much the same manner as that described above. The main differences are that summations over discrete labels become integrations over continuous labels, Kronecker delta symbols become Dirac delta functions, completeness must be assumed (it cannot be proved), and the normalization convention is somewhat different. (See Section 1.15.)

---

[2] We can now appreciate the elegance of Dirac’s notation. The combination of a bra and a ket yields a “bra(c)ket” (which is just a complex number).

1.6 Operators

We have seen that a functional is a machine that inputs a ket vector and spits out a complex number. Consider a somewhat different machine that inputs a ket vector and spits out another ket vector in a deterministic fashion. Mathematicians call such a machine an *operator*. We are only interested in operators that preserve the linear dependencies of the ket vectors upon which they act. Such operators are termed *linear operators*. Consider an operator labeled X. Suppose that when this operator acts on a general ket vector $|A\rangle$ it spits out a new ket vector which is denoted $X\,|A\rangle$. Operator X is linear provided that

$$X\,(|A\rangle + |B\rangle) = X\,|A\rangle + X\,|B\rangle, \tag{1.25}$$

for all ket vectors $|A\rangle$ and $|B\rangle$, and

$$X\,(c\,|A\rangle) = c\,X\,|A\rangle, \tag{1.26}$$

for all complex numbers c. Operators X and Y are said to be equal if

$$X\,|A\rangle = Y\,|A\rangle \tag{1.27}$$

for all kets in the ket space in question. Operator X is termed the *null operator* if

$$X\,|A\rangle = |0\rangle \tag{1.28}$$

for all ket vectors in the space. Operators can be added together. Such addition is defined to obey a commutative and associate algebra: i.e.,

$$X + Y = Y + X, \tag{1.29}$$

$$X + (Y + Z) = (X + Y) + Z. \tag{1.30}$$

Operators can also be multiplied. The multiplication is associative:

$$X\,(Y\,|A\rangle) = (X\,Y)\,|A\rangle = X\,Y\,|A\rangle, \tag{1.31}$$

$$X\,(Y\,Z) = (X\,Y)\,Z = X\,Y\,Z. \tag{1.32}$$

However, in general, it is *non-commutative*: i.e.,

$$X\,Y \neq Y\,X. \tag{1.33}$$

So far, we have only considered linear operators acting on ket vectors. We can also give a meaning to their operation on bra vectors. Consider the inner product of a general bra $\langle B|$ with the ket $X\,|A\rangle$. This product is a number that depends linearly on $|A\rangle$. Thus, it may be considered to be the inner product of $|A\rangle$ with some bra. This bra depends linearly on $\langle B|$, so we may look on it as the result of some linear operator applied to $\langle B|$. This operator is uniquely determined by the original operator X, so we might as well call it the same operator acting on $\langle B|$. A suitable notation to use for the resulting bra when X operates on $\langle B|$ is $\langle B|\,X$. The equation which defines this vector is

$$(\langle B|\,X)\,|A\rangle = \langle B|\,(X\,|A\rangle) \tag{1.34}$$

for any $|A\rangle$ and $\langle B|$. The triple product of $\langle B|$, X, and $|A\rangle$ can be written $\langle B|\,X\,|A\rangle$ without ambiguity, provided we adopt the convention that the bra vector always goes on the left, the operator in the middle, and the ket vector on the right.

Consider the dual bra to $X\,|A\rangle$. This bra depends antilinearly on $|A\rangle$ and must therefore depend linearly on $\langle A|$. Thus, it may be regarded as the result of some linear operator applied to $\langle A|$. This operator is termed the *adjoint* of X, and is denoted $X^\dagger$. Thus,

$$X\,|A\rangle \overset{\text{DC}}{\longleftrightarrow} \langle A|\,X^\dagger. \tag{1.35}$$

It is readily demonstrated that

$$\langle B|\,X^\dagger\,|A\rangle = \langle A|\,X\,|B\rangle^*, \tag{1.36}$$

plus

$$(X\,Y)^\dagger = Y^\dagger\,X^\dagger. \tag{1.37}$$

It is also easily seen that the adjoint of the adjoint of a linear operator is equivalent to the original operator. An *Hermitian* operator ξ has the special property that it is its own adjoint: i.e.,

$$\xi = \xi^\dagger. \tag{1.38}$$

1.7 Outer Product

So far, we have formed the following products: $\langle B|A\rangle$, $X\,|A\rangle$, $\langle A|\,X$, $X\,Y$, $\langle B|\,X\,|A\rangle$. Are there any other products we are allowed to form? How about

$$|B\rangle\langle A|\ ? \tag{1.39}$$

This clearly depends linearly on the bra $\langle A|$ and the ket $|B\rangle$. Suppose that we right-multiply the above product by the general ket $|C\rangle$. We obtain

$$|B\rangle\,\langle A|C\rangle = \langle A|C\rangle\,|B\rangle, \tag{1.40}$$

since $\langle A|C\rangle$ is just a number. Thus, $|B\rangle\langle A|$ acting on a general ket $|C\rangle$ yields another ket. Clearly, the product $|B\rangle\langle A|$ is a linear operator. This operator also acts on bras, as is easily demonstrated by left-multiplying the expression (1.39) by a general bra $\langle C|$. It is also easily demonstrated that

$$(|B\rangle\langle A|)^\dagger = |A\rangle\langle B|. \tag{1.41}$$

Mathematicians term the operator $|B\rangle\langle A|$ the *outer product* of $|B\rangle$ and $\langle A|$. The outer product should not be confused with the inner product, $\langle A|B\rangle$, which is just a number.

1.8 Eigenvalues and Eigenvectors

In general, the ket $X\,|A\rangle$ is not a constant multiple of $|A\rangle$. However, there are some special kets known as the *eigenkets* of operator X. These are denoted

$$|x'\rangle, |x''\rangle, |x'''\rangle, \ \ldots, \tag{1.42}$$

and have the property

$$X\,|x'\rangle = x'\,|x'\rangle, \quad X\,|x''\rangle = x''\,|x''\rangle, \quad \ldots, \tag{1.43}$$

where x', x'', ... are numbers called *eigenvalues*. Clearly, applying X to one of its eigenkets yields the same eigenket multiplied by the associated eigenvalue.

Consider the eigenkets and eigenvalues of a Hermitian operator ξ. These are denoted

$$\xi\,|\xi'\rangle = \xi'\,|\xi'\rangle, \tag{1.44}$$

where $|\xi'\rangle$ is the eigenket associated with the eigenvalue ξ'. Three important results are readily deduced:

(i) The eigenvalues are all real numbers, and the eigenkets corresponding to different eigenvalues are orthogonal. Since ξ is Hermitian, the dual equation to Equation (1.44) (for the eigenvalue ξ'') reads

$$\langle\xi''|\,\xi = \xi''^{\,*}\,\langle\xi''|. \tag{1.45}$$

If we left-multiply Equation (1.44) by $\langle\xi''|$, right-multiply the above equation by $|\xi'\rangle$, and take the difference, we obtain

$$(\xi' - \xi''^{\,*})\,\langle\xi''|\xi'\rangle = 0. \tag{1.46}$$

Suppose that the eigenvalues ξ' and ξ'' are the same. It follows from the above that

$$\xi' = \xi'^{\,*}, \tag{1.47}$$

where we have used the fact that $|\xi'\rangle$ is not the null ket. This proves that the eigenvalues are real numbers. Suppose that the eigenvalues ξ' and ξ'' are different. It follows that

$$\langle\xi''|\xi'\rangle = 0, \tag{1.48}$$

which demonstrates that eigenkets corresponding to different eigenvalues are orthogonal.

(ii) The eigenvalues associated with eigenkets are the same as the eigenvalues associated with eigenbras. An eigenbra of ξ corresponding to an eigenvalue ξ' is defined

$$\langle\xi'|\,\xi = \langle\xi'|\,\xi'. \tag{1.49}$$

(iii) The dual of any eigenket is an eigenbra belonging to the same eigenvalue, and conversely.

1.9 Observables

We have developed a mathematical formalism that comprises three types of objects—bras, kets, and linear operators. We have already seen that kets can be used to represent the possible states of a microscopic system. However, there is a one to one correspondence between the elements of a ket space and its dual bra space, so we must conclude that bras could just as well be used to represent the states of a microscopic system. What about the dynamical variables of the system (e.g., its position, momentum, energy, spin, etc.)? How can these be represented in our formalism? Well, the only objects we have left over are operators. We, therefore, assume that *the dynamical variables of a microscopic system are represented as linear operators acting on the bras and kets that correspond to the various possible states of the system.* Note that the operators have to be linear, otherwise they would, in general, spit out bras/kets pointing in different directions when fed bras/kets pointing in the same direction but differing in length. Since the lengths of bras and kets have no physical significance, it is reasonable to suppose that non-linear operators are also without physical significance.

We have seen that if we observe the polarization state of a photon, by placing a polarizing film in its path then the result is to cause the photon to jump into a state of polarization parallel or perpendicular to the optic axis of the film. The former state is absorbed, and the latter state is transmitted (which is how we tell them apart). In general, we cannot predict into which state a given photon will jump (except in a statistical sense). However, we do know that if the photon is initially polarized parallel to the optic axis then it will definitely be absorbed, and if it is initially polarized perpendicular to the axis then it will definitely be transmitted. We also know that, after passing though the film, a photon must be in a state of polarization perpendicular to the optic axis (otherwise it would not have been transmitted). We can make a second observation of the polarization state of such a photon by placing an identical polarizing film (with the same orientation of the optic axis) immediately behind the first film. It is clear that the photon will definitely be transmitted through the second film.

There is nothing special about the polarization states of a photon. So, more generally, we can say that if a dynamical variable of a microscopic system is measured then the system is caused to jump into one of a number of *independent* states (note that the perpendicular and parallel polarization states of our photon are linearly independent). In general, each of these final states is associated with a different result of the measurement: i.e., a different value of the dynamical variable. Note that the result of the measurement must be a *real* number (there are no measurement machines which output complex numbers). Finally, if an observation is made, and the system is found to be a one particular final state, with one particular value for the dynamical variable, then a second observation, made immediately after the first one, will *definitely* find the system in the same state, and yield the same value for the dynamical variable.

How can we represent all of these facts in our mathematical formalism? Well, by a fairly non-obvious leap of intuition, we are going to assert that *a measurement of a dynamical variable corresponding to an operator X in ket space causes the system to jump into a state corresponding to one of the eigenkets of X.* Not surprisingly, such a state is termed an *eigenstate*. Furthermore, *the result of the measurement is the eigenvalue associated with the eigenket into which the system jumps*. The fact that the result of the measurement must be a real number implies that *dynami-*

cal variables can only be represented by Hermitian operators (because only Hermitian operators are guaranted to have real eigenvalues). The fact that the eigenkets of a Hermitian operator corresponding to different eigenvalues (i.e., different results of the measurement) are orthogonal is in accordance with our earlier requirement that the states into which the system jumps should be mutually independent. We can conclude that the result of a measurement of a dynamical variable represented by a Hermitian operator ξ must be one of the eigenvalues of ξ. Conversely, every eigenvalue of ξ is a possible result of a measurement made on the corresponding dynamical variable. This gives us the physical significance of the eigenvalues. (From now on, the distinction between a state and its representative ket vector, and a dynamical variable and its representative operator, will be dropped, for the sake of simplicity.)

It is reasonable to suppose that *if a certain dynamical variable ξ is measured with the system in a particular state then the states into which the system may jump on account of the measurement are such that the original state is dependent on them.* This fairly innocuous statement has two very important corollaries. First, immediately after an observation whose result is a particular eigenvalue ξ', the system is left in the associated eigenstate. However, this eigenstate is orthogonal to (i.e., independent of) any other eigenstate corresponding to a different eigenvalue. It follows that a second measurement made immediately after the first one must leave the system in an eigenstate corresponding to the eigenvalue ξ'. In other words, the second measurement is bound to give the same result as the first. Furthermore, *if the system is in an eigenstate of ξ, corresponding to an eigenvalue ξ', then a measurement of ξ is bound to give the result ξ'.* This follows because the system cannot jump into an eigenstate corresponding to a different eigenvalue of ξ, because such a state is not dependent on the original state. Second, it stands to reason that a measurement of ξ must always yield some result. It follows that no matter what the initial state of the system, it must always be able to jump into one of the eigenstates of ξ. In other words, a general ket must always be dependent on the eigenkets of ξ. This can only be the case if the eigenkets form a *complete set* (i.e., they span ket space). Thus, *in order for a Hermitian operator ξ to be observable its eigenkets must form a complete set.* A Hermitian operator that satisfies this condition is termed an *observable*. Conversely, any observable quantity must be a Hermitian operator with a complete set of eigenstates.

1.10 Measurements

We have seen that a measurement of some observable ξ of a microscopic system causes the system to jump into one of the eigenstates of ξ. The result of the measurement is the associated eigenvalue (or some function of this quantity). It is impossible to determine into which eigenstate a given system will jump, but it is possible to predict the probability of such a transition. So, what is the probability that a system in some initial state $|A\rangle$ makes a transition to an eigenstate $|\xi'\rangle$ of an observable ξ, as a result of a measurement made on the system? Let us start with the simplest case. If the system is initially in an eigenstate $|\xi'\rangle$ then the transition probability to a eigenstate $|\xi''\rangle$ corresponding to a different eigenvalue is zero, and the transition probability to the same eigenstate $|\xi'\rangle$ is unity. It is convenient to normalize our eigenkets such that they all have unit norms. It follows from the orthogonality property of the eigenkets that

$$\langle\xi'|\xi''\rangle = \delta_{\xi'\xi''}. \tag{1.50}$$

For the moment, we are assuming that the eigenvalues of ξ are all different.

Note that the probability of a transition from an initial eigenstate $|\xi'\rangle$ to a final eigenstate $|\xi''\rangle$ is the same as the value of the inner product $\langle\xi'|\xi''\rangle$. Can we use this correspondence to obtain a general rule for calculating transition probabilities? Well, suppose that the system is initially in a state $|A\rangle$ which is not an eigenstate of ξ. Can we identify the transition probability to a final eigenstate $|\xi'\rangle$ with the inner product $\langle A|\xi'\rangle$? In fact, we cannot because $\langle A|\xi'\rangle$ is, in general, a complex number, and complex probabilities do not make any sense. Let us try again. Suppose that we identify the transition probability with the *modulus squared* of the inner product, $|\langle A|\xi'\rangle|^2$? This quantity is definitely a positive number (so it could be a probability). This guess also gives the right answer for the transition probabilities between eigenstates. In fact, it is the correct guess.

Because the eigenstates of an observable ξ form a complete set, we can express any given state $|A\rangle$ as a linear combination of them. It is easily demonstrated that

$$|A\rangle = \sum_{\xi'} |\xi'\rangle\langle\xi'|A\rangle, \tag{1.51}$$

$$\langle A| = \sum_{\xi'} \langle A|\xi'\rangle\langle\xi'|, \tag{1.52}$$

$$\langle A|A\rangle = \sum_{\xi'} \langle A|\xi'\rangle\langle\xi'|A\rangle = \sum_{\xi'} |\langle A|\xi'\rangle|^2, \tag{1.53}$$

where the summation is over all the different eigenvalues of ξ, and use has been made of Equation (1.20), as well as the fact that the eigenstates are mutually orthogonal. Note that all of the above results follow from the extremely useful (and easily proved) result

$$\sum_{\xi'} |\xi'\rangle\langle\xi'| = 1, \tag{1.54}$$

where 1 denotes the identity operator. The relative probability of a transition to an eigenstate $|\xi'\rangle$, which is equivalent to the relative probability of a measurement of ξ yielding the result ξ', is

$$P(\xi') \propto |\langle A|\xi'\rangle|^2. \tag{1.55}$$

The absolute probability is clearly

$$P(\xi') = \frac{|\langle A|\xi'\rangle|^2}{\sum_{\xi'} |\langle A|\xi'\rangle|^2} = \frac{|\langle A|\xi'\rangle|^2}{\langle A|A\rangle}. \tag{1.56}$$

If the ket $|A\rangle$ is normalized such that its norm is unity then this probability simply reduces to

$$P(\xi') = |\langle A|\xi'\rangle|^2. \tag{1.57}$$

1.11 Expectation Values

Consider an *ensemble* of microscopic systems prepared in the same initial state $|A\rangle$. Suppose that a measurement of the observable ξ is made on each system. We know that each measurement yields the value ξ' with probability $P(\xi')$. What is the mean value of the measurement? This quantity, which is generally referred to as the *expectation value* of ξ, is given by

$$\langle\xi\rangle = \sum_{\xi'} \xi' P(\xi') = \sum_{\xi'} \xi' \, |\langle A|\xi'\rangle|^2 = \sum_{\xi'} \xi' \langle A|\xi'\rangle\langle\xi'|A\rangle = \sum_{\xi'} \langle A|\,\xi\,|\xi'\rangle\langle\xi'|A\rangle, \tag{1.58}$$

which reduces to

$$\langle\xi\rangle = \langle A|\,\xi\,|A\rangle \tag{1.59}$$

with the aid of Equation (1.54).

Consider the identity operator, 1. All states are eigenstates of this operator with the eigenvalue unity. Thus, the expectation value of this operator is always unity: i.e.,

$$\langle A|\,1\,|A\rangle = \langle A|A\rangle = 1, \tag{1.60}$$

for all $|A\rangle$. Note that it is only possible to normalize a given ket $|A\rangle$, such that Equation (1.60) is satisfied, because of the more general property (1.21) of the norm. This property depends on the previously adopted correspondence (1.16) between the elements of a ket space and those of its dual bra space.

1.12 Degeneracy

Suppose that two different eigenstates $|\xi_a'\rangle$ and $|\xi_b'\rangle$ of the observable ξ correspond to the *same* eigenvalue ξ'. These states are termed *degenerate* eigenstates. Degenerate eigenstates are necessarily orthogonal to any eigenstates corresponding to different eigenvalues, but, in general, they are not orthogonal to each other (i.e., the proof of orthogonality given in Section 1.8 does not work in this case). This is unfortunate, because much of the previous formalism depends crucially on the mutual orthogonality of the different eigenstates of an observable. Note, however, that any linear combination of $|\xi_a'\rangle$ and $|\xi_b'\rangle$ is also an eigenstate corresponding to the eigenvalue ξ'. It follows that we can always *construct* two mutually orthogonal degenerate eigenstates. For instance,

$$|\xi_1'\rangle = |\xi_a'\rangle, \tag{1.61}$$

$$|\xi_2'\rangle = \frac{|\xi_b'\rangle - \langle\xi_a'|\xi_b'\rangle|\xi_a'\rangle}{[1 - |\langle\xi_a'|\xi_b'\rangle|^2]^{1/2}}. \tag{1.62}$$

This result is easily generalized to the case of more than two degenerate eigenstates. We conclude that it is always possible to construct a complete set of mutually orthogonal eigenstates for any given observable.

1.13 Compatible Observables

Suppose that we wish to simultaneously measure two observables, ξ and η, of a microscopic system? Let us assume that we possess an apparatus that is capable of measuring ξ, and another that can measure η. For instance, the two observables in question might be the projection in the x- and z-directions of the spin angular momentum of a spin one-half particle. These could be measured using appropriate Stern-Gerlach apparatuses (see Sakurai, Section 1.1). Suppose that we make a measurement of ξ, and the system is consequently thrown into one of the eigenstates of ξ, $|\xi'\rangle$, with eigenvalue ξ'. What happens if we now make a measurement of η? Well, suppose that the eigenstate $|\xi'\rangle$ is also an eigenstate of η, with eigenvalue η'. In this case, a measurement of η will definitely give the result η'. A second measurement of ξ will definitely give the result ξ', and so on. In this sense, we can say that the observables ξ and η *simultaneously* have the values ξ' and η', respectively. Clearly, if all eigenstates of ξ are also eigenstates of η then it is always possible to make a simultaneous measurement of ξ and η. Such observables are termed *compatible*.

Suppose, however, that the eigenstates of ξ are not eigenstates of η. Is it still possible to measure both observables simultaneously? Let us again make an observation of ξ which throws the system into an eigenstate $|\xi'\rangle$, with eigenvalue ξ'. We can now make a second observation to determine η. This will throw the system into one of the (many) eigenstates of η which depend on $|\xi'\rangle$. In principle, each of these eigenstates is associated with a different result of the measurement. Suppose that the system is thrown into an eigenstate $|\eta'\rangle$, with the eigenvalue η'. Another measurement of ξ will throw the system into one of the (many) eigenstates of ξ which depend on $|\eta'\rangle$. Each eigenstate is again associated with a different possible result of the measurement. It is clear that if the observables ξ and η do not possess simultaneous eigenstates then if the value of ξ is known (i.e., the system is in an eigenstate of ξ) then the value of η is uncertain (i.e., the system is not in an eigenstate of η), and vice versa. We say that the two observables are *incompatible*.

We have seen that *the condition for two observables ξ and η to be simultaneously measurable is that they should possess simultaneous eigenstates* (i.e., every eigenstate of ξ should also be an eigenstate of η). Suppose that this is the case. Let a general eigenstate of ξ, with eigenvalue ξ', also be an eigenstate of η, with eigenvalue η'. It is convenient to denote this simultaneous eigenstate $|\xi'\eta'\rangle$. We have

$$\xi\,|\xi'\eta'\rangle = \xi'\,|\xi'\eta'\rangle, \tag{1.63}$$

$$\eta\,|\xi'\eta'\rangle = \eta'\,|\xi'\eta'\rangle. \tag{1.64}$$

We can left-multiply the first equation by η, and the second equation by ξ, and then take the difference. The result is

$$(\xi\,\eta - \eta\,\xi)\,|\xi'\eta'\rangle = |0\rangle \tag{1.65}$$

for each simultaneous eigenstate. Recall that the eigenstates of an observable must form a complete set. It follows that the simultaneous eigenstates of two observables must also form a complete set. Thus, the above equation implies that

$$(\xi\,\eta - \eta\,\xi)\,|A\rangle = |0\rangle, \tag{1.66}$$

where $|A\rangle$ is a general ket. The only way that this can be true is if

$$\xi\,\eta = \eta\,\xi. \tag{1.67}$$

Thus, *the condition for two observables ξ and η to be simultaneously measurable is that they should commute*.

1.14 Uncertainty Relation

We have seen that if ξ and η are two noncommuting observables then a determination of the value of ξ leaves the value of η uncertain, and vice versa. It is possible to quantify this uncertainty. For a general observable ξ, we can define a Hermitian operator

$$\Delta\xi = \xi - \langle\xi\rangle, \tag{1.68}$$

where the expectation value is taken over the particular physical state under consideration. It is obvious that the expectation value of $\Delta\xi$ is zero. The expectation value of $(\Delta\xi)^2 \equiv \Delta\xi\,\Delta\xi$ is termed the *variance* of ξ, and is, in general, non-zero. In fact, it is easily demonstrated that

$$\langle(\Delta\xi)^2\rangle = \langle\xi^2\rangle - \langle\xi\rangle^2. \tag{1.69}$$

The variance of ξ is a measure of the uncertainty in the value of ξ for the particular state in question (i.e., it is a measure of the width of the distribution of likely values of ξ about the expectation value). If the variance is zero then there is no uncertainty, and a measurement of ξ is bound to give the expectation value, $\langle\xi\rangle$.

Consider the Schwarz inequality

$$\langle A|A\rangle\langle B|B\rangle \geq |\langle A|B\rangle|^2, \tag{1.70}$$

which is analogous to

$$|\mathbf{a}|^2\,|\mathbf{b}|^2 \geq |\mathbf{a}\cdot\mathbf{b}|^2 \tag{1.71}$$

in Euclidian space. This inequality can be proved by noting that

$$(\langle A| + c^*\,\langle B|)\,(|A\rangle + c\,|B\rangle) \geq 0, \tag{1.72}$$

where c is any complex number. If c takes the special value $-\langle B|A\rangle/\langle B|B\rangle$ then the above inequality reduces to

$$\langle A|A\rangle\langle B|B\rangle - |\langle A|B\rangle|^2 \geq 0, \tag{1.73}$$

which is the same as the Schwarz inequality.

Let us substitute

$$|A\rangle = \Delta\xi\,|\,\rangle, \tag{1.74}$$

$$|B\rangle = \Delta\eta\,|\,\rangle, \tag{1.75}$$

into the Schwarz inequality, where the blank ket $|\,\rangle$ stands for any general ket. We find

$$\langle(\Delta\xi)^2\rangle\langle(\Delta\eta)^2\rangle \geq |\langle\Delta\xi\,\Delta\eta\rangle|^2, \tag{1.76}$$

where use has been made of the fact that $\Delta\xi$ and $\Delta\eta$ are Hermitian operators. Note that

$$\Delta\xi\,\Delta\eta = \frac{1}{2}\,[\Delta\xi, \Delta\eta] + \frac{1}{2}\,\{\Delta\xi, \Delta\eta\}, \tag{1.77}$$

where the *commutator*, $[\Delta\xi, \Delta\eta]$, and the *anti-commutator*, $\{\Delta\xi, \Delta\eta\}$, are defined

$$[\Delta\xi, \Delta\eta] \equiv \Delta\xi\,\Delta\eta - \Delta\eta\,\Delta\xi, \tag{1.78}$$

$$\{\Delta\xi, \Delta\eta\} \equiv \Delta\xi\,\Delta\eta + \Delta\eta\,\Delta\xi. \tag{1.79}$$

The commutator is clearly *anti-Hermitian*,

$$([\Delta\xi, \Delta\eta])^\dagger = (\Delta\xi\,\Delta\eta - \Delta\eta\,\Delta\xi)^\dagger = \Delta\eta\,\Delta\xi - \Delta\xi\,\Delta\eta = -\,[\Delta\xi, \Delta\eta]\,, \tag{1.80}$$

whereas the anti-commutator is obviously Hermitian. Now, it is easily demonstrated that the expectation value of an Hermitian operator is a real number, whereas the expectation value of an anti-Hermitian operator is an imaginary number. It follows that the right-hand side of

$$\langle\Delta\xi\,\Delta\eta\rangle = \frac{1}{2}\,\langle[\Delta\xi, \Delta\eta]\rangle + \frac{1}{2}\,\langle\{\Delta\xi, \Delta\eta\}\rangle, \tag{1.81}$$

consists of the sum of an imaginary and a real number. Taking the modulus squared of both sides gives

$$|\langle\Delta\xi\,\Delta\eta\rangle|^2 = \frac{1}{4}\,|\langle[\xi, \eta]\rangle|^2 + \frac{1}{4}\,|\langle\{\Delta\xi, \Delta\eta\}\rangle|^2, \tag{1.82}$$

where use has been made of $\langle\Delta\xi\rangle = 0$, etc. The final term on the right-hand side of the above expression is positive definite, so we can write

$$\langle(\Delta\xi)^2\rangle\,\langle(\Delta\eta)^2\rangle \geq |\langle\Delta\xi\,\Delta\eta\rangle|^2 \geq \frac{1}{4}\,|\langle[\xi, \eta]\rangle|^2, \tag{1.83}$$

where use has been made of Equation (1.76). The above expression is termed the *uncertainty relation*. According to this relation, an exact knowledge of the value of ξ implies no knowledge whatsoever of the value of η, and vice versa. The one exception to this rule is when ξ and η commute, in which case exact knowledge of ξ does not necessarily imply no knowledge of η.

1.15 Continuous Spectra

Up to now, we have studiously avoided dealing with observables possessing eigenvalues that lie in a continuous range, rather than having discrete values. The reason for this is that continuous eigenvalues imply a ket space of nondenumerably infinite dimension. Unfortunately, continuous eigenvalues are unavoidable in quantum mechanics. In fact, the most important observables of all—namely position and momentum—generally have continuous eigenvalues. Fortunately, many of the results that we obtained previously for a finite-dimensional ket space with discrete eigenvalues can be generalized to ket spaces of nondenumerably infinite dimensions.

Suppose that ξ is an observable with continuous eigenvalues. We can still write the eigenvalue equation as

$$\xi\,|\xi'\rangle = \xi'\,|\xi'\rangle. \tag{1.84}$$

But, ξ' can now take a continuous range of values. Let us assume, for the sake of simplicity, that ξ' can take any value. The orthogonality condition (1.50) generalizes to

$$\langle\xi'|\xi''\rangle = \delta(\xi'-\xi''), \tag{1.85}$$

where $\delta(x)$ denotes the famous *Dirac delta function*, and satisfies

$$\int_{-\infty}^{\infty} dx\, f(x)\,\delta(x-x') = f(x') \tag{1.86}$$

for any function, $f(x)$, that is well-behaved at $x = x'$. Note that there are clearly a nondenumerably infinite number of mutually orthogonal eigenstates of ξ. Hence, the dimensionality of ket space is nondenumerably infinite. Furthermore, eigenstates corresponding to a continuous range of eigenvalues *cannot* be normalized such that they have unit norms. In fact, these eigenstates have *infinite* norms: i.e., they are infinitely long. This is the major difference between eigenstates in a finite-dimensional and an infinite-dimensional ket space. The extremely useful relation (1.54) generalizes to

$$\int d\xi'\,|\xi'\rangle\langle\xi'| = 1. \tag{1.87}$$

Note that a summation over discrete eigenvalues goes over into an integral over a continuous range of eigenvalues. The eigenstates $|\xi'\rangle$ must form a complete set if ξ is to be an observable. It follows that any general ket can be expanded in terms of the $|\xi'\rangle$. In fact, the expansions (1.51)–(1.53) generalize to

$$|A\rangle = \int d\xi'\,|\xi'\rangle\langle\xi'|A\rangle, \tag{1.88}$$

$$\langle A| = \int d\xi'\,\langle A|\xi'\rangle\langle\xi'|, \tag{1.89}$$

$$\langle A|A\rangle = \int d\xi'\,\langle A|\xi'\rangle\langle\xi'|A\rangle = \int d\xi'\,|\langle A|\xi'\rangle|^2. \tag{1.90}$$

These results also follow simply from Equation (1.87). We have seen that it is not possible to normalize the eigenstates $|\xi'\rangle$ such that they have unit norms. Fortunately, this convenient normalization is still possible for a general state vector. In fact, according to Equation (1.90), the normalization condition can be written

$$\langle A|A\rangle = \int d\xi'\,|\langle A|\xi'\rangle|^2 = 1. \tag{1.91}$$

We have now studied observables whose eigenvalues take a discrete number of values, as well as those whose eigenvalues take a continuous range of values. There are a number of other cases we could look at. For instance, observables whose eigenvalues can take a (finite) continuous range of values, plus a set of discrete values. Such cases can be dealt with using a fairly straightforward generalization of the previous analysis (see Dirac, Chapters II and III).

Exercises

1.1 According to classical physics, a non-relativistic electron whose instantaneous acceleration is of magnitude a radiates electromagnetic energy at the rate

$$P = \frac{e^2\,a^2}{6\pi\,\epsilon_0\,c^3},$$

where e is the magnitude of the electron charge, ϵ_0 the permittivity of the vacuum, and m_e the electron mass. Consider a classical electron in a circular orbit of radius r around a proton. Demonstrate that the radiated energy would cause the orbital radius to decrease in time according to

$$\frac{d}{dt}\left(\frac{r}{a_0}\right)^3 = -\frac{1}{\tau},$$

where $a_0 = 4\pi\,\epsilon_0\,\hbar^2/(m_e\,e^2)$ is the Bohr radius, $\hbar$ the reduced Planck constant, and

$$\tau = \frac{a_0}{4\,\alpha^4\,c}.$$

Here, c is the velocity of light in a vacuum, and $\alpha = e^2/(4\pi\,\epsilon_0\,\hbar\,c)$ the fine structure constant. Deduce that the classical lifetime of a hydrogen atom is $\tau \simeq 1.6\times 10^{-11}$ s.

1.2 Demonstrate that

$$\langle B|A\rangle = \langle A|B\rangle^*$$

in a finite dimensional ket space.

1.3 Demonstrate that in a finite dimensional ket space:

(a)

$$\langle B|\,X^\dagger\,|A\rangle = \langle A|\,X\,|B\rangle^*.$$

(b)

$$(X\,Y)^\dagger = Y^\dagger\,X^\dagger.$$

(c)

$$(X^\dagger)^\dagger = X.$$

(d)

$$(|B\rangle\langle A|)^\dagger = |A\rangle\langle B|.$$

Here, X, Y are general operators.

1.4 If A, B are Hermitian operators then demonstrate that $A\,B$ is only Hermitian provided A and B commute. In addition, show that $(A+B)^n$ is Hermitian, where n is a positive integer.

1.5 Let A be a general operator. Show that $A+A^\dagger$, $\mathrm{i}\,(A-A^\dagger)$, and $A\,A^\dagger$ are Hermitian operators.

1.6 Let H be an Hermitian operator. Demonstrate that the Hermitian conjugate of the operator $\exp(\mathrm{i}\,H) \equiv \sum_{n=0,\infty}(\mathrm{i}\,H)^n/n!$ is $\exp(-\mathrm{i}\,H)$.

1.7 Let the $|\xi'\rangle$ be the eigenkets of an observable ξ, whose corresponding eigenvalues, ξ', are discrete. Demonstrate that

$$\sum_{\xi'} |\xi'\rangle\langle\xi'| = 1,$$

where the sum is over all eigenvalues, and 1 denotes the unity operator.

1.8 Let the $|\xi_i'\rangle$, where $i = 1, N$, and $N > 1$, be a set of degenerate eigenkets of some observable ξ. Suppose that the $|\xi_i'\rangle$ are not mutually orthogonal. Demonstrate that a set of mutually orthogonal (but unnormalized) degenerate eigenkets, $|\xi_i''\rangle$, for $i = 1, N$, can be constructed as follows:

$$|\xi_i''\rangle = |\xi_i'\rangle - \sum_{j=1,i-1} \frac{\langle\xi_j''|\xi_i'\rangle}{\langle\xi_j''|\xi_j''\rangle}\,|\xi_j''\rangle.$$

This process is known as *Gram-Schmidt orthogonalization*.

1.9 Demonstrate that the expectation value of a Hermitian operator is a real number. Show that the expectation value of an anti-hermitian operator is an imaginary number.

1.10 Let H be an Hermitian operator. Demonstrate that $\langle H^2\rangle \geq 0$.

1.11 Consider an Hermitian operator, H, that has the property that $H^4 = 1$, where 1 is the unity operator. What are the eigenvalues of H? What are the eigenvalues if H is not restricted to being Hermitian?

1.12 Let ξ be an observable whose eigenvalues, ξ', lie in a continuous range. Let the $|\xi'\rangle$, where

$$\langle\xi'|\xi''\rangle = \delta(\xi' - \xi''),$$

be the corresponding eigenkets. Demonstrate that

$$\int d\xi'\, |\xi'\rangle\langle\xi'| = 1,$$

where the integral is over the whole range of eigenvalues, and 1 denotes the unity operator.

2 Position and Momentum

2.1 Introduction

So far, we have considered *general* dynamical variables represented by *general* linear operators acting in ket space. However, in classical mechanics, the most important dynamical variables are those involving position and momentum. Let us investigate the role of such variables in quantum mechanics.

In classical mechanics, the position q and momentum p of some component of a dynamical system are represented as *real numbers* which, by definition, commute. In quantum mechanics, these quantities are represented as *non-commuting* linear Hermitian operators acting in a ket space that represents all of the possible states of the system. Our first task is to discover a quantum mechanical replacement for the classical result $q\,p - p\,q = 0$.

2.2 Poisson Brackets

Consider a dynamical system whose state at a particular time t is fully specified by N independent classical coordinates q_i (where i runs from 1 to N). Associated with each generalized coordinate q_i is a classical canonical momentum p_i. For instance, a Cartesian coordinate has an associated linear momentum, an angular coordinate has an associated angular momentum, etc. As is well-known, the behavior of a classical system can be specified in terms of Lagrangian or Hamiltonian dynamics. For instance, in Hamiltonian dynamics,

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \tag{2.1}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \tag{2.2}$$

where the function $H(q_i, p_i, t)$ is the system energy at time t expressed in terms of the classical coordinates and canonical momenta. This function is usually referred to as the *Hamiltonian* of the system.

We are interested in finding some construct in classical dynamics that consists of *products* of dynamical variables. If such a construct exists then we hope to generalize it somehow to obtain a rule describing how dynamical variables commute with one another in quantum mechanics. There is, indeed, one well-known construct in classical dynamics that involves products of dynamical variables. The classical *Poisson bracket* of two dynamical variables u and v is defined

$$[u, v]_{cl} = \sum_{i=1,N}\left(\frac{\partial u}{\partial q_i}\frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i}\frac{\partial v}{\partial q_i}\right), \tag{2.3}$$

where u and v are regarded as functions of the coordinates and momenta, q_i and p_i. It is easily demonstrated that

$$[q_i, q_j]_{cl} = 0, \tag{2.4}$$

$$[p_i, p_j]_{cl} = 0, \tag{2.5}$$

$$[q_i, p_j]_{cl} = \delta_{ij}. \tag{2.6}$$

The time evolution of a dynamical variable can also be written in terms of a Poisson bracket by noting that

$$\frac{du}{dt} = \sum_{i=1,N}\left(\frac{\partial u}{\partial q_i}\frac{dq_i}{dt} + \frac{\partial u}{\partial p_i}\frac{dp_i}{dt}\right) = \sum_{i=1,N}\left(\frac{\partial u}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial u}{\partial p_i}\frac{\partial H}{\partial q_i}\right) = [u, H]_{cl}, \tag{2.7}$$

where use has been made of Hamilton's equations, (2.1)–(2.2).

Can we construct a quantum mechanical Poisson bracket in which u and v are non-commuting operators, instead of functions? Well, the main properties of the classical Poisson bracket are as follows:

$$[u, v]_{cl} = -[v, u]_{cl}, \tag{2.8}$$

$$[u, c]_{cl} = 0, \tag{2.9}$$

$$[u_1 + u_2, v]_{cl} = [u_1, v]_{cl} + [u_2, v]_{cl}, \tag{2.10}$$

$$[u, v_1 + v_2]_{cl} = [u, v_1]_{cl} + [u, v_2]_{cl}, \tag{2.11}$$

$$[u_1\, u_2, v]_{cl} = [u_1, v]_{cl}\, u_2 + u_1\, [u_2, v]_{cl}, \tag{2.12}$$

$$[u, v_1\, v_2]_{cl} = [u, v_1]_{cl}\, v_2 + v_1\, [u, v_2]_{cl}, \tag{2.13}$$

$$[u, [v, w]_{cl}]_{cl} + [v, [w, u]_{cl}]_{cl} + [w, [u, v]_{cl}]_{cl} = 0. \tag{2.14}$$

The last relation is known as the *Jacobi identity*. In the above, u, v, w, etc., represent dynamical variables, and c represents a number. Can we find some combination of non-commuting operators u and v, etc., that satisfies all of the above relations? We shall refer to such a combination as a quantum mechanical Poisson bracket.

Actually, we can evaluate the quantum mechanical Poisson bracket $[u_1\, u_2, v_1\, v_2]_{qm}$ in two different ways, because we can employ either of the formulae (2.12) or (2.13) first. Thus,

$$\begin{aligned}[u_1\, u_2, v_1\, v_2]_{qm} &= [u_1, v_1\, v_2]_{qm}\, u_2 + u_1\, [u_2, v_1\, v_2]_{qm}\\ &= \Big([u_1, v_1]_{qm}\, v_2 + v_1\, [u_1, v_2]_{qm}\Big)\, u_2 + u_1\, \Big([u_2, v_1]_{qm}\, v_2 + v_1\, [u_2, v_2]_{qm}\Big)\\ &= [u_1, v_1]_{qm}\, v_2\, u_2 + v_1\, [u_1, v_2]_{qm}\, u_2 + u_1\, [u_2, v_1]_{qm}\, v_2 + u_1\, v_1\, [u_2, v_2]_{qm},\end{aligned} \tag{2.15}$$

and

$$\begin{aligned}[u_1\, u_2, v_1\, v_2]_{qm} &= [u_1\, u_2, v_1]_{qm}\, v_2 + v_1\, [u_1\, u_2, v_2]_{qm}\\ &= [u_1, v_1]_{qm}\, u_2\, v_2 + u_1\, [u_2, v_1]_{qm}\, v_2 + v_1\, [u_1, v_2]_{qm}\, u_2 + v_1\, u_1\, [u_2, v_2]_{qm}.\end{aligned} \tag{2.16}$$

Note that the order of the various factors has been preserved, because they now represent *non-commuting* operators. Equating the above two results yields

$$[u_1, v_1]_{qm}\,(u_2\,v_2 - v_2\,u_2) = (u_1\,v_1 - v_1\,u_1)\,[u_2, v_2]_{qm}. \tag{2.17}$$

Since this relation must hold for u_1 and v_1 quite independent of u_2 and v_2, it follows that

$$u_1\,v_1 - v_1\,u_1 = \mathrm{i}\,\hbar\,[u_1, v_1]_{qm}, \tag{2.18}$$

$$u_2\,v_2 - v_2\,u_2 = \mathrm{i}\,\hbar\,[u_2, v_2]_{qm}, \tag{2.19}$$

where $\hbar$ does not depend on u_1, v_1, u_2, v_2, and also commutes with $(u_1\,v_1 - v_1\,u_1)$. Because u_1, etc., are general operators, it follows that $\hbar$ is just a number. We want the quantum mechanical Poisson bracket of two Hermitian operators to be a Hermitian operator itself, because the classical Poisson bracket of two real dynamical variables is real. This requirement is satisfied if $\hbar$ is a real number. Thus, the quantum mechanical Poisson bracket of two dynamical variables u and v is given by

$$[u, v]_{qm} = \frac{u\,v - v\,u}{\mathrm{i}\,\hbar}, \tag{2.20}$$

where $\hbar$ is a new universal constant of nature. Quantum mechanics agrees with experiments provided that $\hbar$ takes the value $h/2\pi$, where

$$h = 6.6261 \times 10^{-34}\ \ \mathrm{J\,s} \tag{2.21}$$

is *Planck's constant*. The notation $[u, v]$ is conventionally reserved for the commutator $u\,v - v\,u$ in quantum mechanics. Thus,

$$[u, v]_{qm} = \frac{[u, v]}{\mathrm{i}\,\hbar}. \tag{2.22}$$

It is easily demonstrated that the quantum mechanical Poisson bracket, as defined above, satisfies all of the relations (2.8)–(2.14).

The strong analogy we have found between the classical Poisson bracket, defined in Equation (2.3), and the quantum mechanical Poisson bracket, defined in Equation (2.22), leads us to assume that the quantum mechanical bracket has the same value as the corresponding classical bracket, at least for the simplest cases. In other words, we are assuming that Equations (2.4)–(2.6) hold for quantum mechanical as well as classical Poisson brackets. This argument yields the fundamental commutation relations

$$[q_i, q_j] = 0, \tag{2.23}$$

$$[p_i, p_j] = 0, \tag{2.24}$$

$$[q_i, p_j] = \mathrm{i}\,\hbar\,\delta_{ij}. \tag{2.25}$$

These results provide us with the basis for calculating commutation relations between general dynamical variables. For instance, if two dynamical variables, ξ and η, can both be written as a power series in the q_i and p_i then repeated application of Equations (2.8)–(2.13) allows $[\xi, \eta]$ to be expressed in terms of the fundamental commutation relations (2.23)–(2.25).

Equations (2.23)–(2.25) provide the foundation for the analogy between quantum mechanics and classical mechanics. Note that the classical result (that everything commutes) is obtained in the limit $\hbar \rightarrow 0$. Thus, *classical mechanics can be regarded as the limiting case of quantum mechanics when $\hbar$ goes to zero.* In classical mechanics, each pair of generalized coordinate and its conjugate momentum, q_i and p_i, correspond to a different classical degree of freedom of the system. It is clear from Equations (2.23)–(2.25) that in quantum mechanics the *dynamical variables corresponding to different degrees of freedom all commute*. It is only those variables corresponding to the same degree of freedom that may fail to commute.

2.3 Wavefunctions

Consider a simple system with one classical degree of freedom, which corresponds to the Cartesian coordinate x. Suppose that x is free to take any value (e.g., x could be the position of a free particle). The classical dynamical variable x is represented in quantum mechanics as a linear Hermitian operator which is also called x. Moreover, the operator x possesses eigenvalues x' lying in the *continuous* range $-\infty < x' < +\infty$ (since the eigenvalues correspond to all the possible results of a measurement of x). We can span ket space using the suitably normalized eigenkets of x. An eigenket corresponding to the eigenvalue x' is denoted $|x'\rangle$. Moreover, [see Equation (1.85)]

$$\langle x'|x''\rangle = \delta(x' - x''). \tag{2.26}$$

The eigenkets satisfy the extremely useful relation [see Equation (1.87)]

$$\int_{-\infty}^{+\infty} dx'\, |x'\rangle\langle x'| = 1. \tag{2.27}$$

This formula expresses the fact that the eigenkets are complete, mutually orthogonal, and suitably normalized.

A state ket $|A\rangle$ (which represents a general state A of the system) can be expressed as a linear superposition of the eigenkets of the position operator using Equation (2.27). Thus,

$$|A\rangle = \int_{-\infty}^{+\infty} dx'\, \langle x'|A\rangle |x'\rangle \tag{2.28}$$

The quantity $\langle x'|A\rangle$ is a complex function of the position eigenvalue x'. We can write

$$\langle x'|A\rangle = \psi_A(x'). \tag{2.29}$$

Here, $\psi_A(x')$ is the famous *wavefunction* of quantum mechanics. Note that state A is completely specified by its wavefunction $\psi_A(x')$ [because the wavefunction can be used to reconstruct the state ket $|A\rangle$ using Equation (2.28)]. It is clear that the wavefunction of state A is simply the collection of the weights of the corresponding state ket $|A\rangle$, when it is expanded in terms of the eigenkets of the position operator. Recall, from Section 1.10, that the probability of a measurement of a dynamical variable ξ yielding the result ξ' when the system is in state A is given by $|\langle \xi'|A\rangle|^2$, assuming that the eigenvalues of ξ are discrete. This result is easily generalized to dynamical variables possessing continuous eigenvalues. In fact, the probability of a measurement of x yielding a result lying in the range x' to $x' + dx'$ when the system is in a state $|A\rangle$ is $|\langle x'|A\rangle|^2\, dx'$. In other words, the probability of a measurement of position yielding a result in the range x' to $x' + dx'$ when the wavefunction of the system is $\psi_A(x')$ is

$$P(x', dx') = |\psi_A(x')|^2\, dx'. \tag{2.30}$$

This formula is only valid if the state ket $|A\rangle$ is properly normalized: i.e., if $\langle A|A\rangle = 1$. The corresponding normalization for the wavefunction is

$$\int_{-\infty}^{+\infty} dx'\, |\psi_A(x')|^2 = 1. \tag{2.31}$$

Consider a second state B represented by a state ket $|B\rangle$ and a wavefunction $\psi_B(x')$. The inner product $\langle B|A\rangle$ can be written

$$\langle B|A\rangle = \int_{-\infty}^{+\infty} dx'\, \langle B|x'\rangle\langle x'|A\rangle = \int_{-\infty}^{+\infty} dx'\, \psi_B^*(x')\, \psi_A'(x'), \tag{2.32}$$

where use has been made of Equations (2.27) and (2.29). Thus, the inner product of two states is related to the overlap integral of their wavefunctions.

Consider a general function $f(x)$ of the observable x [e.g., $f(x) = x^2$]. If $|B\rangle = f(x)\,|A\rangle$ then it follows that

$$\psi_B(x') = \langle x'|f(x) \int_{-\infty}^{+\infty} dx''\, \psi_A(x'')\, |x''\rangle = \int_{-\infty}^{+\infty} dx''\, f(x'')\, \psi_A(x'')\, \langle x'|x''\rangle, \tag{2.33}$$

giving

$$\psi_B(x') = f(x')\, \psi_A(x'), \tag{2.34}$$

where use has been made of Equation (2.26). Here, $f(x')$ is the same function of the position eigenvalue x' that $f(x)$ is of the position operator x: i.e., if $f(x) = x^2$ then $f(x') = x'^2$. It follows, from the above result, that a general state ket $|A\rangle$ can be written

$$|A\rangle = \psi_A(x)\rangle, \tag{2.35}$$

where $\psi_A(x)$ is the same function of the operator x that the wavefunction $\psi_A(x')$ is of the position eigenvalue x', and the ket $\rangle$ has the wavefunction $\psi(x') = 1$. The ket $\rangle$ is termed the *standard ket*. The dual of the standard ket is termed the *standard bra*, and is denoted $\langle$. It is easily seen that

$$\langle \psi_A^*(x) \xleftrightarrow{\text{DC}} \psi_A(x)\rangle. \tag{2.36}$$

Note, finally, that $\psi_A(x)\rangle$ is often shortened to $\psi_A\rangle$, leaving the dependence on the position operator x tacitly understood.

2.4 Schrödinger Representation

Consider the simple system described in the previous section. A general state ket can be written $\psi(x)\rangle$, where $\psi(x)$ is a general function of the position operator x, and $\psi(x')$ is the associated wavefunction. Consider the ket whose wavefunction is $d\psi(x')/dx'$. This ket is denoted $d\psi/dx\rangle$. The new ket is clearly a linear function of the original ket, so we can think of it as the result of some linear operator acting on $\psi\rangle$. Let us denote this operator d/dx. It follows that

$$\frac{d}{dx}\,\psi\rangle = \frac{d\psi}{dx}\rangle. \tag{2.37}$$

Any linear operator that acts on ket vectors can also act on bra vectors. Consider d/dx acting on a general bra $\langle\phi(x)$. According to Equation (1.34), the bra $\langle\phi\,d/dx$ satisfies

$$\left(\langle\phi\,\frac{d}{dx}\right)\psi\rangle = \langle\phi\left(\frac{d}{dx}\,\psi\rangle\right). \tag{2.38}$$

Making use of Equations (2.27) and (2.29), we can write

$$\int_{-\infty}^{+\infty} dx'\,\langle\phi\,\frac{d}{dx}\,|x'\rangle\,\psi(x') = \int_{-\infty}^{+\infty} dx'\,\phi(x')\,\frac{d\psi(x')}{dx'}. \tag{2.39}$$

The right-hand side can be transformed via integration by parts to give

$$\int_{-\infty}^{+\infty} dx'\,\langle\phi\,\frac{d}{dx}\,|x'\rangle\,\psi(x') = -\int_{-\infty}^{+\infty} dx'\,\frac{d\phi(x')}{dx'}\,\psi(x'), \tag{2.40}$$

assuming that the contributions from the limits of integration vanish. It follows that

$$\langle\phi\,\frac{d}{dx}\,|x'\rangle = -\frac{d\phi(x')}{dx'}, \tag{2.41}$$

which implies that

$$\langle\phi\,\frac{d}{dx} = -\langle\frac{d\phi}{dx}. \tag{2.42}$$

The neglect of contributions from the limits of integration in Equation (2.40) is reasonable because physical wavefunctions are square-integrable [see Equation (2.31)]. Note that

$$\frac{d}{dx}\,\psi\rangle = \frac{d\psi}{dx}\rangle \overset{\text{DC}}{\longleftrightarrow} \langle\frac{d\psi^*}{dx} = -\langle\psi^*\,\frac{d}{dx}, \tag{2.43}$$

where use has been made of Equation (2.42). It follows, by comparison with Equations (1.35) and (2.36), that

$$\left(\frac{d}{dx}\right)^\dagger = -\frac{d}{dx}. \tag{2.44}$$

Thus, d/dx is an *anti-Hermitian* operator.

Let us evaluate the commutation relation between the operators x and d/dx. We have

$$\frac{d}{dx}\,x\,\psi\rangle = \frac{d(x\,\psi)}{dx}\rangle = x\,\frac{d}{dx}\,\psi\rangle + \psi\rangle. \tag{2.45}$$

Since this holds for any ket $\psi\rangle$, it follows that

$$\frac{d}{dx}\,x - x\,\frac{d}{dx} = 1. \tag{2.46}$$

Let p_x be the momentum conjugate to x (for the simple system under consideration p_x is a straightforward linear momentum). According to Equation (2.25), x and p satisfy the commutation relation

$$x\,p_x - p_x\,x = \mathrm{i}\,\hbar. \tag{2.47}$$

It can be seen, by comparison with Equation (2.46), that the Hermitian operator $-\mathrm{i}\,\hbar\,d/dx$ *satisfies the same commutation relation with x that p_x does*. The most general conclusion which may be drawn from a comparison of Equations (2.46) and (2.47) is that

$$p_x = -\mathrm{i}\,\hbar\,\frac{d}{dx} + f(x), \tag{2.48}$$

since (as is easily demonstrated) a general function $f(x)$ of the position operator automatically commutes with x.

We have chosen to normalize the eigenkets and eigenbras of the position operator such that they satisfy the normalization condition (2.26). However, this choice of normalization does not uniquely determine the eigenkets and eigenbras. Suppose that we transform to a new set of eigenbras which are related to the old set via

$$\langle x'|_{\rm new} = \mathrm{e}^{\,\mathrm{i}\,\gamma'}\langle x'|_{\rm old}, \tag{2.49}$$

where $\gamma' \equiv \gamma(x')$ is a real function of x'. This transformation amounts to a rearrangement of the relative phases of the eigenbras. The new normalization condition is

$$\begin{aligned}\langle x'|x''\rangle_{\rm new} &= \langle x'|\,\mathrm{e}^{\,\mathrm{i}\,\gamma'}\mathrm{e}^{-\mathrm{i}\,\gamma''}|x''\rangle_{\rm old} = \mathrm{e}^{\,\mathrm{i}\,(\gamma'-\gamma'')}\langle x'|x''\rangle_{\rm old}\\ &= \mathrm{e}^{\,\mathrm{i}\,(\gamma'-\gamma'')}\,\delta(x'-x'') = \delta(x'-x''). \end{aligned}\tag{2.50}$$

Thus, the new eigenbras satisfy the same normalization condition as the old eigenbras.

By definition, the standard ket $\rangle$ satisfies $\langle x'|\rangle = 1$. It follows from Equation (2.49) that the new standard ket is related to the old standard ket via

$$\rangle_{\rm new} = \mathrm{e}^{-\mathrm{i}\,\gamma}\rangle_{\rm old}, \tag{2.51}$$

where $\gamma \equiv \gamma(x)$ is a real function of the position operator x. The dual of the above equation yields the transformation rule for the standard bra,

$$\langle_{\rm new} = \langle_{\rm old}\,\mathrm{e}^{\,\mathrm{i}\,\gamma}. \tag{2.52}$$

The transformation rule for a general operator A follows from Equations (2.51) and (2.52), plus the requirement that the triple product $\langle A\rangle$ remain invariant (this must be the case, otherwise the probability of a measurement yielding a certain result would depend on the choice of eigenbras). Thus,

$$A_{\rm new} = {\rm e}^{-{\rm i}\,\gamma} A_{\rm old}\,{\rm e}^{\,{\rm i}\,\gamma}. \tag{2.53}$$

Of course, if A commutes with x then A is invariant under the transformation. In fact, d/dx is the only operator (that we know of) which does not commute with x, so Equation (2.53) yields

$$\left(\frac{d}{dx}\right)_{\rm new} = {\rm e}^{-{\rm i}\,\gamma}\frac{d}{dx}\,{\rm e}^{\,{\rm i}\,\gamma} = \frac{d}{dx} + {\rm i}\,\frac{d\gamma}{dx}, \tag{2.54}$$

where the subscript "old" is taken as read. It follows, from Equation (2.48), that the momentum operator p_x can be written

$$p_x = -{\rm i}\,\hbar\left(\frac{d}{dx}\right)_{\rm new} - \hbar\,\frac{d\gamma}{dx} + f(x). \tag{2.55}$$

Thus, the special choice

$$\hbar\,\gamma(x) = \int^x dx'\,f(x') \tag{2.56}$$

yields

$$p_x = -{\rm i}\,\hbar\left(\frac{d}{dx}\right)_{\rm new}. \tag{2.57}$$

Equation (2.56) fixes γ to within an arbitrary additive constant: i.e., the special eigenkets and eigenbras for which Equation (2.57) is true are determined to within an arbitrary common phase-factor.

In conclusion, it is possible to find a set of basis eigenkets and eigenbras of the position operator x that satisfy the normalization condition (2.26), and for which the momentum conjugate to x can be represented as the operator

$$p_x = -{\rm i}\,\hbar\,\frac{d}{dx}. \tag{2.58}$$

A general state ket is written $\psi(x)\rangle$, where the standard ket $\rangle$ satisfies $\langle x'|\rangle = 1$, and where $\psi(x') = \langle x'|\psi(x)\rangle$ is the wavefunction. This scheme of things is known as the *Schrödinger representation*, and is the basis of wave mechanics.

2.5 Generalized Schrödinger Representation

In the preceding section, we developed the Schrödinger representation for the case of a single operator x corresponding to a classical Cartesian coordinate. However, this scheme can easily be extended. Consider a system with N generalized coordinates, $q_1\cdots q_N$, which can all be *simultaneously* measured. These are represented as N *commuting* operators, $q_1\cdots q_N$, each with a continuous range of eigenvalues, $q_1'\cdots q_N'$. Ket space is conveniently spanned by the simultaneous eigenkets of $q_1\cdots q_N$, which are denoted $|q_1'\cdots q_N'\rangle$. These eigenkets must form a complete set, otherwise the $q_1\cdots q_N$ would not be simultaneously observable.

The orthogonality condition for the eigenkets [i.e., the generalization of Equation (2.26)] is

$$\langle q_1' \cdots q_N' | q_1'' \cdots q_N'' \rangle = \delta(q_1' - q_1'')\,\delta(q_2' - q_2'') \cdots \delta(q_N' - q_N''). \tag{2.59}$$

The completeness condition [i.e., the generalization of Equation (2.27)] is

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} dq_1' \cdots dq_N' \, | q_1' \cdots q_N' \rangle \langle q_1' \cdots q_N' | = 1. \tag{2.60}$$

The standard ket $\rangle$ is defined such that

$$\langle q_1' \cdots q_N' | \rangle = 1. \tag{2.61}$$

The standard bra $\langle$ is the dual of the standard ket. A general state ket is written

$$\psi(q_1 \cdots q_N)\rangle. \tag{2.62}$$

The associated wavefunction is

$$\psi(q_1' \cdots q_N') = \langle q_1' \cdots q_N' | \psi \rangle. \tag{2.63}$$

Likewise, a general state bra is written

$$\langle \phi(q_1 \cdots q_N), \tag{2.64}$$

where

$$\phi(q_1' \cdots q_N') = \langle \phi | q_1' \cdots q_N' \rangle. \tag{2.65}$$

The probability of an observation of the system simultaneously finding the first coordinate in the range q_1' to $q_1' + dq_1'$, the second coordinate in the range q_2' to $q_2' + dq_2'$, etc., is

$$P(q_1' \cdots q_N'; dq_1' \cdots dq_N') = |\psi(q_1' \cdots q_N')|^2 \, dq_1' \cdots dq_N'. \tag{2.66}$$

Finally, the normalization condition for a physical wavefunction is

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} dq_1' \cdots dq_N' \, |\psi(q_1' \cdots q_N')|^2 = 1. \tag{2.67}$$

The N linear operators $\partial/\partial q_i$ (where i runs from 1 to N) are defined

$$\frac{\partial}{\partial q_i}\,\psi\rangle = \frac{\partial \psi}{\partial q_i}\rangle. \tag{2.68}$$

These linear operators can also act on bras (provided the associated wavefunctions are square integrable) in accordance with [see Equation (2.42)]

$$\langle \phi \, \frac{\partial}{\partial q_i} = -\langle \frac{\partial \phi}{\partial q_i}. \tag{2.69}$$

Corresponding to Equation (2.46), we can derive the commutation relations

$$\frac{\partial}{\partial q_i}\,q_j - q_j\,\frac{\partial}{\partial q_i} = \delta_{ij}. \tag{2.70}$$

It is also clear that

$$\frac{\partial}{\partial q_i}\,\frac{\partial}{\partial q_j}\,\psi\rangle = \frac{\partial^2\psi}{\partial q_i\partial q_j}\rangle = \frac{\partial}{\partial q_j}\,\frac{\partial}{\partial q_i}\,\psi\rangle, \tag{2.71}$$

showing that

$$\frac{\partial}{\partial q_i}\,\frac{\partial}{\partial q_j} = \frac{\partial}{\partial q_j}\,\frac{\partial}{\partial q_i}. \tag{2.72}$$

It can be seen, by comparison with Equations (2.23)–(2.25), that the linear operators $-\mathrm{i}\,\hbar\,\partial/\partial q_i$ satisfy the same commutation relations with the q's and with each other that the p's do. The most general conclusion we can draw from this coincidence of commutation relations is (see Dirac)

$$p_i = -\mathrm{i}\,\hbar\frac{\partial}{\partial q_i} + \frac{\partial F(q_1\cdots q_N)}{\partial q_i}. \tag{2.73}$$

However, the function F can be transformed away via a suitable readjustment of the phases of the basis eigenkets (see Section 2.4, and Dirac). Thus, we can always construct a set of simultaneous eigenkets of $q_1\cdots q_N$ for which

$$p_i = -\mathrm{i}\,\hbar\frac{\partial}{\partial q_i}. \tag{2.74}$$

This is the generalized Schrödinger representation.

It follows from Equations (2.61), (2.68), and (2.74) that

$$p_i\rangle = 0. \tag{2.75}$$

Thus, the standard ket in the Schrödinger representation is a simultaneous eigenket of all the momentum operators belonging to the eigenvalue zero. Note that

$$\langle q_1'\cdots q_N'|\,\frac{\partial}{\partial q_i}\,\psi\rangle = \langle q_1'\cdots q_N'|\,\frac{\partial\psi}{\partial q_i}\rangle = \frac{\partial\psi(q_1'\cdots q_N')}{\partial q_i'} = \frac{\partial}{\partial q_i'}\,\langle q_1'\cdots q_N'|\psi\rangle. \tag{2.76}$$

Hence,

$$\langle q_1'\cdots q_N'|\,\frac{\partial}{\partial q_i} = \frac{\partial}{\partial q_i'}\,\langle q_1'\cdots q_N'|, \tag{2.77}$$

so that

$$\langle q_1'\cdots q_N'|\,p_i = -\mathrm{i}\,\hbar\,\frac{\partial}{\partial q_i'}\,\langle q_1'\cdots q_N'|. \tag{2.78}$$

The dual of the above equation gives

$$p_i\,|q_1'\cdots q_N'\rangle = \mathrm{i}\,\hbar\,\frac{\partial}{\partial q_i'}\,|q_1'\cdots q_N'\rangle. \tag{2.79}$$

2.6 Momentum Representation

Consider a system with one degree of freedom, describable in terms of a coordinate x and its conjugate momentum p_x, both of which have a continuous range of eigenvalues. We have seen that it is possible to represent the system in terms of the eigenkets of x. This is termed the Schrödinger representation. However, it is also possible to represent the system in terms of the eigenkets of p_x.

Consider the eigenkets of p_x which belong to the eigenvalues p_x'. These are denoted $|p_x'\rangle$. The orthogonality relation for the momentum eigenkets is

$$\langle p_x'|p_x''\rangle = \delta(p_x' - p_x''), \tag{2.80}$$

and the corresponding completeness relation is

$$\int_{-\infty}^{+\infty} dp_x'\, |p_x'\rangle\langle p_x'| = 1. \tag{2.81}$$

A general state ket can be written

$$\phi(p)\rangle \tag{2.82}$$

where the standard ket $\rangle$ satisfies

$$\langle p_x'|\rangle = 1. \tag{2.83}$$

Note that the standard ket in this representation is quite different to that in the Schrödinger representation. The momentum space wavefunction $\phi(p_x')$ satisfies

$$\phi(p_x') = \langle p_x'|\phi\rangle. \tag{2.84}$$

The probability that a measurement of the momentum yields a result lying in the range p_x' to $p_x' + dp_x'$ is given by

$$P(p_x', dp_x') = |\phi(p_x')|^2\, dp_x'. \tag{2.85}$$

Finally, the normalization condition for a physical momentum space wavefunction is

$$\int_{-\infty}^{+\infty} dp_x'\, |\phi(p_x')|^2 = 1. \tag{2.86}$$

The fundamental commutation relations (2.23)–(2.25) exhibit a particular symmetry between coordinates and their conjugate momenta. If all the coordinates are transformed into their conjugate momenta, and vice versa, and i is then replaced by −i, then the commutation relations are unchanged. It follows from this symmetry that we can always choose the eigenkets of p_x in such a manner that the coordinate x can be represented as (see Section 2.4)

$$x = \mathrm{i}\,\hbar\,\frac{d}{dp_x}. \tag{2.87}$$

This is termed the momentum representation.

The above result is easily generalized to a system with more than one degree of freedom. Suppose the system is specified by N coordinates, $q_1 \cdots q_N$, and N conjugate momenta, $p_1 \cdots p_N$. Then, in the momentum representation, the coordinates can be written as

$$q_i = \mathrm{i}\,\hbar\,\frac{\partial}{\partial p_i}. \tag{2.88}$$

We also have

$$q_i\rangle = 0, \tag{2.89}$$

and

$$\langle p_1' \cdots p_N'|\,q_i = \mathrm{i}\,\hbar\,\frac{\partial}{\partial p_i'}\,\langle p_1' \cdots p_N'|. \tag{2.90}$$

The momentum representation is less useful than the Schrödinger representation for a very simple reason. The energy operator (i.e., the Hamiltonian) of most simple systems takes the form of a sum of quadratic terms in the momenta (i.e., the kinetic energy) plus a complicated function of the coordinates (i.e., the potential energy). In the Schrödinger representation, the eigenvalue problem for the energy translates into a second-order differential equation in the coordinates, with a complicated potential function. In the momentum representation, the problem transforms into a high-order differential equation in the momenta, with a quadratic potential. With the mathematical tools at our disposal, we are far better able to solve the former type of problem than the latter. Hence, the Schrödinger representation is generally more useful than the momentum representation.

2.7 Uncertainty Relation

How is a momentum space wavefunction related to the corresponding coordinate space wavefunction? To answer this question, let us consider the representative $\langle x'|p_x'\rangle$ of the momentum eigenkets $|p'\rangle$ in the Schrödinger representation for a system with a single degree of freedom. This representative satisfies

$$p_x'\,\langle x'|p_x'\rangle = \langle x'|\,p_x\,|p_x'\rangle = -\mathrm{i}\,\hbar\,\frac{d}{dx'}\,\langle x'|p_x'\rangle, \tag{2.91}$$

where use has been made of Equation (2.78) (for the case of a system with one degree of freedom). The solution of the above differential equation is

$$\langle x'|p_x'\rangle = c'\,\exp(\,\mathrm{i}\,p_x'\,x'/\hbar), \tag{2.92}$$

where $c' = c'(p_x')$. It is easily demonstrated that

$$\langle p_x'|p_x''\rangle = \int_{-\infty}^{+\infty} dx'\,\langle p_x'|x'\rangle\,\langle x'|p_x''\rangle = c'^{\,*}\,c''\int_{-\infty}^{\infty} dx'\,\exp[-\mathrm{i}\,(p_x' - p_x'')\,x'/\hbar]. \tag{2.93}$$

The well-known mathematical result

$$\int_{-\infty}^{+\infty} dx\,\exp(\,\mathrm{i}\,a\,x) = 2\pi\,\delta(a), \tag{2.94}$$

yields

$$\langle p_x'|p_x''\rangle = |c'|^2\, h\, \delta(p_x' - p_x''). \tag{2.95}$$

This is consistent with Equation (2.80), provided that $c' = h^{-1/2}$. Thus,

$$\langle x'|p_x'\rangle = h^{-1/2}\, \exp(\,\mathrm{i}\, p_x'\, x'/\hbar). \tag{2.96}$$

Consider a general state ket $|A\rangle$ whose coordinate wavefunction is $\psi(x')$, and whose momentum wavefunction is $\Psi(p_x')$. In other words,

$$\psi(x') = \langle x'|A\rangle, \tag{2.97}$$

$$\Psi(p_x') = \langle p_x'|A\rangle. \tag{2.98}$$

It is easily demonstrated that

$$\psi(x') = \int_{-\infty}^{+\infty} dp_x'\, \langle x'|p_x'\rangle\langle p_x'|A\rangle = \frac{1}{h^{1/2}}\int_{-\infty}^{+\infty} dp_x'\, \Psi(p_x')\, \exp(\,\mathrm{i}\, p_x'\, x'/\hbar) \tag{2.99}$$

and

$$\Psi(p_x') = \int_{-\infty}^{+\infty} dx'\, \langle p_x'|x'\rangle\langle x'|A\rangle = \frac{1}{h^{1/2}}\int_{-\infty}^{+\infty} dx'\, \psi(x')\, \exp(-\mathrm{i}\, p_x' x'/\hbar), \tag{2.100}$$

where use has been made of Equations (2.27), (2.81), (2.94), and (2.96). Clearly, the momentum space wavefunction is the *Fourier transform* of the coordinate space wavefunction.

Consider a state whose coordinate space wavefunction is a *wavepacket*. In other words, the wavefunction only has non-negligible amplitude in some spatially localized region of extent $\varDelta x$. As is well-known, the Fourier transform of a wavepacket fills up a wavenumber band of approximate extent $\varDelta k \sim 1/\varDelta x$. Note that in Equation (2.99) the role of the wavenumber k is played by the quantity $p_x'/\hbar$. It follows that the momentum space wavefunction corresponding to a wavepacket in coordinate space extends over a range of momenta $\varDelta p_x \sim \hbar/\varDelta x$. Clearly, a measurement of x is almost certain to give a result lying in a range of width $\varDelta x$. Likewise, measurement of p_x is almost certain to yield a result lying in a range of width $\varDelta p_x$. The product of these two uncertainties is

$$\varDelta x\, \varDelta p_x \sim \hbar. \tag{2.101}$$

This result is called the *Heisenberg uncertainty principle*.

Actually, it is possible to write the Heisenberg uncertainty principle more exactly by making use of Equation (1.83) and the commutation relation (2.47). We obtain

$$\langle(\varDelta x)^2\rangle\, \langle(\varDelta p_x)^2\rangle \geq \frac{\hbar^2}{4} \tag{2.102}$$

for any general state. It is easily demonstrated that the minimum uncertainty states, for which the equality sign holds in the above relation, correspond to Gaussian wavepackets in both coordinate and momentum space.

2.8 Displacement Operators

Consider a system with one degree of freedom corresponding to the Cartesian coordinate x. Suppose that we displace this system some distance along the x-axis. We could imagine that the system is on wheels, and we just give it a little push. The final state of the system is completely determined by its initial state, together with the direction and magnitude of the displacement. Note that the type of displacement we are considering is one in which *everything* to do with the system is displaced. So, if the system is subject to an external potential then the potential must be displaced.

The situation is not so clear with state kets. The final state of the system only determines the *direction* of the displaced state ket. Even if we adopt the convention that all state kets have unit norms, the final ket is still not completely determined, because it can be multiplied by a constant phase-factor. However, we know that the superposition relations between states remain invariant under the displacement. This follows because the superposition relations have a physical significance that is unaffected by a displacement of the system. Thus, if

$$|R\rangle = |A\rangle + |B\rangle \tag{2.103}$$

in the undisplaced system, and the displacement causes ket $|R\rangle$ to transform to ket $|Rd\rangle$, etc., then in the displaced system we have

$$|Rd\rangle = |Ad\rangle + |Bd\rangle. \tag{2.104}$$

Incidentally, this determines the displaced kets to within a single arbitrary phase-factor to be multiplied into all of them. The displaced kets cannot be multiplied by individual phase-factors, because this would wreck the superposition relations.

Since Equation (2.104) holds in the displaced system whenever Equation (2.103) holds in the undisplaced system, it follows that the displaced ket $|Rd\rangle$ must be the result of some linear operator acting on the undisplaced ket $|R\rangle$. In other words,

$$|Rd\rangle = D\,|R\rangle, \tag{2.105}$$

where D an operator that depends only on the nature of the displacement. The arbitrary phase-factor by which all displaced kets may be multiplied results in D being undetermined to an arbitrary multiplicative constant of modulus unity.

We now adopt the ansatz that any combination of bras, kets, and dynamical variables that possesses a physical significance is invariant under a displacement of the system. The normalization condition

$$\langle A|A\rangle = 1 \tag{2.106}$$

for a state ket $|A\rangle$ certainly has a physical significance. Thus, we must have

$$\langle Ad|Ad\rangle = 1. \tag{2.107}$$

Now, $|Ad\rangle = D\,|A\rangle$ and $\langle Ad| = \langle A|\,D^\dagger$, so

$$\langle A|\,D^\dagger\,D\,|A\rangle = 1. \tag{2.108}$$

Because this must hold for any state ket $|A\rangle$, it follows that

$$D^\dagger D = 1. \tag{2.109}$$

Hence, the displacement operator is *unitary*. Note that the above relation implies that

$$|A\rangle = D^\dagger |Ad\rangle. \tag{2.110}$$

The equation

$$v\,|A\rangle = |B\rangle, \tag{2.111}$$

where the operator v represents a dynamical variable, has some physical significance. Thus, we require that

$$v_d\,|Ad\rangle = |Bd\rangle, \tag{2.112}$$

where v_d is the displaced operator. It follows that

$$v_d\,|Ad\rangle = D\,|B\rangle = D\,v\,|A\rangle = D\,v\,D^\dagger\,|Ad\rangle. \tag{2.113}$$

Since this is true for any ket $|Ad\rangle$, we have

$$v_d = D\,v\,D^\dagger. \tag{2.114}$$

Note that the arbitrary numerical factor in D does not affect either of the results (2.109) and (2.114).

Suppose, now, that the system is displaced an *infinitesimal* distance δx along the x-axis. We expect that the displaced ket $|Ad\rangle$ should approach the undisplaced ket $|A\rangle$ in the limit as $\delta x \to 0$. Thus, we expect the limit

$$\lim_{\delta x\to 0}\frac{|Ad\rangle - |A\rangle}{\delta x} = \lim_{\delta x\to 0}\frac{D-1}{\delta x}\,|A\rangle \tag{2.115}$$

to exist. Let

$$d_x = \lim_{\delta x\to 0}\frac{D-1}{\delta x}, \tag{2.116}$$

where d_x is denoted the *displacement operator* along the x-axis. The fact that D can be replaced by $D\,\exp(\,{\rm i}\,\gamma)$, where γ is a real phase-angle, implies that d_x can be replaced by

$$\lim_{\delta x\to 0}\frac{D\,\exp({\rm i}\,\gamma)-1}{\delta x} = \lim_{\delta x\to 0}\frac{D-1+{\rm i}\,\gamma}{\delta x} = d_x + {\rm i}\,a_x, \tag{2.117}$$

where a_x is the limit of $\gamma/\delta x$. We have assumed, as seems reasonable, that γ tends to zero as $\delta x \to 0$. It is clear that the displacement operator is undetermined to an arbitrary imaginary additive constant.

For small δx, we have

$$D = 1 + \delta x\,d_x. \tag{2.118}$$

It follows from Equation (2.109) that

$$(1 + \delta x\,d_x^\dagger)\,(1 + \delta x\,d_x) = 1. \tag{2.119}$$

Neglecting order $(\delta x)^2$, we obtain

$$d_x^\dagger + d_x = 0. \tag{2.120}$$

Thus, the displacement operator is *anti-Hermitian*. Substituting into Equation (2.114), and again neglecting order $(\delta x)^2$, we find that

$$v_d = (1 + \delta x\, d_x)\, v\, (1 - \delta x\, d_x) = v + \delta x\, (d_x\, v - v\, d_x), \tag{2.121}$$

which implies

$$\lim_{\delta x\to 0} \frac{v_d - v}{\delta x} = d_x\, v - v\, d_x. \tag{2.122}$$

Let us consider a specific example. Suppose that a state has a wavefunction $\psi(x')$. If the system is displaced a distance δx along the x-axis then the new wavefunction is $\psi(x' - \delta x)$ (i.e., the same shape shifted in the x-direction by a distance δx). Actually, the new wavefunction can be multiplied by an arbitrary number of modulus unity. It can be seen that the new wavefunction is obtained from the old wavefunction according to the prescription $x' \to x' - \delta x$. Thus,

$$x_d = x - \delta x. \tag{2.123}$$

A comparison with Equation (2.122), using $x = v$, yields

$$d_x\, x - x\, d_x = -1. \tag{2.124}$$

It follows that $\mathrm{i}\,\hbar\, d_x$ obeys the same commutation relation with x that p_x, the momentum conjugate to x, does [see Equation (2.25)]. The most general conclusion we can draw from this observation is that

$$p_x = \mathrm{i}\,\hbar\, d_x + f(x), \tag{2.125}$$

where f is Hermitian (since p_x is Hermitian). However, the fact that d_x is undetermined to an arbitrary additive imaginary constant (which could be a function of x) enables us to transform the function f out of the above equation, leaving

$$p_x = \mathrm{i}\,\hbar\, d_x. \tag{2.126}$$

Thus, the displacement operator in the x-direction is proportional to the momentum conjugate to x. We say that p_x is the *generator* of translations along the x-axis.

A finite translation along the x-axis can be constructed from a series of very many infinitesimal translations. Thus, the operator $D(\Delta x)$ which translates the system a distance Δx along the x-axis is written

$$D(\Delta x) = \lim_{N\to\infty} \left(1 - \mathrm{i}\,\frac{\Delta x}{N}\frac{p_x}{\hbar}\right)^N, \tag{2.127}$$

where use has been made of Equations (2.118) and (2.126). It follows that

$$D(\Delta x) = \exp\left(-\mathrm{i}\, p_x\, \Delta x/\hbar\right). \tag{2.128}$$

The unitary nature of the operator is now clearly apparent.

We can also construct displacement operators which translate the system along the y- and z-axes. Note that a displacement a distance Δx along the x-axis *commutes* with a displacement a distance Δy along the y-axis. In other words, if the system is moved Δx along the x-axis, and then Δy along the y-axis, then it ends up in the same state as if it were moved Δy along the y-axis, and then Δx along the x-axis. The fact that translations in independent directions commute is clearly associated with the fact that the conjugate momentum operators associated with these directions also commute [see Equations (2.24) and (2.128)].

Exercises

2.1 Demonstrate that

$$[q_i, q_j]_{cl} = 0,$$
$$[p_i, p_j]_{cl} = 0,$$
$$[q_i, p_j]_{cl} = \delta_{ij},$$

where $[\cdots, \cdots]_{cl}$ represents a classical Poisson bracket. Here, the q_i and p_i are the coordinates and corresponding canonical momenta of a classical, many degree of freedom, dynamical system.

2.2 Verify that

$$\begin{aligned}
[u, v] &= -[v, u],\\
[u, c] &= 0,\\
[u_1 + u_2, v] &= [u_1, v] + [u_2, v],\\
[u, v_1 + v_2] &= [u, v_1] + [u, v_2],\\
[u_1\, u_2, v] &= [u_1, v]\, u_2 + u_1\, [u_2, v],\\
[u, v_1\, v_2] &= [u, v_1]\, v_2 + v_1\, [u, v_2], [u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0,
\end{aligned}$$

where $[\cdots, \cdots]$ represents either a classical or a quantum mechanical Poisson bracket. Here, u, u, w, etc., represent dynamical variables (i.e., functions of the coordinates and canonical momenta of a dynamical system), and c represents a number.

2.3 Consider a Gaussian wavepacket whose corresponding wavefunction is

$$\psi(x') = \psi_0 \exp\left[-\frac{(x' - x_0)^2}{4\,\sigma^2}\right],$$

where ψ_0, x_0, and σ are constants. Demonstrate that

(a)

$$\langle x\rangle = x_0,$$

(b)

$$\langle(\Delta x)^2\rangle = \sigma^2,$$

(c)

$$\langle p\rangle = 0,$$

(d)

$$\langle(\Delta p)^2\rangle = \frac{\hbar^2}{4\,\sigma^2}.$$

Here, x and p are a position operator and its conjugate momentum operator, respectively.

2.4 Suppose that we displace a one-dimensional quantum mechanical system a distance a along the x-axis. The corresponding displacement operator is

$$D(a) = \exp\left(-\mathrm{i}\,p_x\,a/\hbar\right),$$

where p_x is the momentum conjugate to the position operator x. Demonstrate that

$$D(a)\,x\,D(a)^\dagger = x - a.$$

[Hint: Use the momentum representation, $x = \mathrm{i}\,\hbar\,d/dp_x$.] Similarly, demonstrate that

$$D(a)\,x^m\,D(a)^\dagger = (x-a)^m.$$

Hence, deduce that

$$D(a)\,V(x)\,D(a)^\dagger = V(x-a),$$

where $V(x)$ is a general function of x.

Let $k = p_x/\hbar$, and let $|k'\rangle$ denote an eigenket of the k operator belonging to the eigenvalue k'. Demonstrate that

$$|A\rangle = \sum_{n=-\infty,\infty} c_n\,|k' + n\,k_a\rangle,$$

where the c_n are arbitrary complex coefficients, and $k_a = 2\pi/a$, is an eigenket of the $D(a)$ operator belonging to the eigenvalue $\exp(-\mathrm{i}\,k'\,a)$. Show that the corresponding wavefunction can be written

$$\psi_A(x') = \mathrm{e}^{\,\mathrm{i}\,k'\,x'}\,u(x'),$$

where $u(x'+a) = u(x')$ for all x'.

3 Quantum Dynamics

3.1 Schrödinger Equation of Motion

Up to now, we have only considered systems at one particular instant of time. Let us now investigate the time evolution if quantum mechanical systems.

Consider a system in a state A that evolves in time. At time t, the state of the system is represented by the ket $|At\rangle$. The label A is needed to distinguish this ket from any other ket ($|Bt\rangle$, say) that is evolving in time. The label t is needed to distinguish the different states of the system at different times.

The final state of the system at time t is completely determined by its initial state at time t_0 plus the time interval $t - t_0$ (assuming that the system is left undisturbed during this time interval). However, the final state only determines the *direction* of the final state ket. Even if we adopt the convention that all state kets have unit norms, the final ket is still not completely determined, because it can be multiplied by an arbitrary phase-factor. However, we expect that if a superposition relation holds for certain states at time t_0 then the same relation should hold between the corresponding time-evolved states at time t, assuming that the system is left undisturbed between times t_0 and t. In other words, if

$$|Rt_0\rangle = |At_0\rangle + |Bt_0\rangle \tag{3.1}$$

for any three kets then we should have

$$|Rt\rangle = |At\rangle + |Bt\rangle. \tag{3.2}$$

This rule determines the time-evolved kets to within a single arbitrary phase-factor to be multiplied into all of them. The evolved kets cannot be multiplied by individual phase-factors because this would invalidate the superposition relation at later times.

According to Equations (3.1) and (3.2), the final ket $|Rt\rangle$ depends linearly on the initial ket $|Rt_0\rangle$. Thus, the final ket can be regarded as the result of some linear operator acting on the initial ket: i.e.,

$$|Rt\rangle = T\,|Rt_0\rangle, \tag{3.3}$$

where T is a linear operator that depends only on the times t and t_0. The arbitrary phase-factor by which all time-evolved kets may be multiplied results in $T(t, t_0)$ being undetermined to an arbitrary multiplicative constant of modulus unity.

Because we have adopted a convention in which the norm of any state ket is unity, it make sense to define the time evolution operator T in such a manner that it preserves the length of any ket upon which it acts (i.e., if a ket is properly normalized at time t then it will remain normalized at all subsequent times $t > t_0$). This is always possible, because the length of a ket possesses no physical significance. Thus, we require that

$$\langle At_0|At_0\rangle = \langle At|At\rangle \tag{3.4}$$

for any ket A, which immediately yields

$$T^\dagger\, T = 1. \tag{3.5}$$

Hence, the time evolution operator T is *unitary*.

Up to now, the time evolution operator T looks very much like the spatial displacement operator D introduced in the previous section. However, there are some important differences between time evolution and spatial displacement. In general, we *do* expect the expectation value of some observable ξ to evolve with time, even if the system is left in a state of undisturbed motion (after all, time evolution has no meaning unless something *observable* changes with time). The triple product $\langle A|\,\xi\,|A\rangle$ can evolve either because the ket $|A\rangle$ evolves and the operator ξ stays constant, the ket $|A\rangle$ stays constant and the operator ξ evolves, or both the ket $|A\rangle$ and the operator ξ evolve. Because we are already committed to evolving state kets, according to Equation (3.3), let us assume that the time evolution operator T can be chosen in such a manner that the operators representing the dynamical variables of the system *do not* evolve in time (unless they contain some specific time dependence).

We expect, from physical continuity, that if $t \rightarrow t_0$ then $|At\rangle \rightarrow |At_0\rangle$ for any ket A. Thus, the limit

$$\lim_{t\rightarrow t_0} \frac{|At\rangle - |At_0\rangle}{t - t_0} = \lim_{t\rightarrow t_0} \frac{T - 1}{t - t_0}\, |At_0\rangle \tag{3.6}$$

should exist. Note that this limit is simply the derivative of $|At_0\rangle$ with respect to t_0. Let

$$\tau(t_0) = \lim_{t\rightarrow t_0} \frac{T(t, t_0) - 1}{t - t_0}. \tag{3.7}$$

It is easily demonstrated from Equation (3.5) that τ is anti-Hermitian: i.e.,

$$\tau^\dagger + \tau = 0. \tag{3.8}$$

The fact that T can be replaced by $T \exp(\,{\rm i}\,\gamma)$ (where γ is real) implies that τ is undetermined to an arbitrary *imaginary* additive constant (see Section 2.8). Let us define the Hermitian operator $H(t_0) = {\rm i}\,\hbar\,\tau$. This operator is undetermined to an arbitrary *real* additive constant. It follows from Equations (3.6) and (3.7) that

$${\rm i}\,\hbar\,\frac{d|At_0\rangle}{dt_0} = {\rm i}\,\hbar \lim_{t\rightarrow t_0} \frac{|At\rangle - |At_0\rangle}{t - t_0} = {\rm i}\,\hbar\,\tau(t_0)\,|At_0\rangle = H(t_0)\,|At_0\rangle. \tag{3.9}$$

When written for general t, this equation becomes

$${\rm i}\,\hbar\,\frac{d|At\rangle}{dt} = H(t)\,|At\rangle. \tag{3.10}$$

Equation (3.10) gives the general law for the time evolution of a state ket in a scheme in which the operators representing the dynamical variables remain fixed. This equation is denoted the *Schrödinger equation of motion*. It involves a Hermitian operator $H(t)$ which is, presumably, a characteristic of the dynamical system under investigation.

We saw, in Section 2.8, that if the operator $D(x, x_0)$ displaces the system along the x-axis from x_0 to x then

$$p_x = \mathrm{i}\,\hbar \lim_{x\rightarrow x_0} \frac{D(x, x_0) - 1}{x - x_0}, \tag{3.11}$$

where p_x is the operator representing the momentum conjugate to x. Furthermore, we have just shown that if the operator $T(t, t_0)$ evolves the system in time from t_0 to t then

$$H(t_0) = \mathrm{i}\,\hbar \lim_{t\rightarrow t_0} \frac{T(t, t_0) - 1}{t - t_0}. \tag{3.12}$$

Thus, the dynamical variable corresponding to the operator H stands to time t as the momentum p_x stands to the coordinate x. By analogy with classical physics, this suggests that $H(t)$ is the operator representing the total *energy* of the system. (Recall that, in classical physics, if the equations of motion of a system are invariant under an x-displacement of the system then this implies that the system conserves momentum in the x-direction. Likewise, if the equations of motion are invariant under a temporal displacement then this implies that the system conserves energy.) The operator $H(t)$ is usually called the *Hamiltonian* of the system. The fact that the Hamiltonian is undetermined to an arbitrary real additive constant is related to the well-known phenomenon that energy is undetermined to an arbitrary additive constant in physics (i.e., the zero of potential energy is not well-defined).

Substituting $|At\rangle = T\,|At_0\rangle$ into Equation (3.10) yields

$$\mathrm{i}\,\hbar\,\frac{dT}{dt}\,|At_0\rangle = H(t)\,T\,|At_0\rangle. \tag{3.13}$$

Because this must hold for any initial state $|At_0\rangle$, we conclude that

$$\mathrm{i}\,\hbar\,\frac{dT}{dt} = H(t)\,T. \tag{3.14}$$

This equation can be integrated to give

$$T(t, t_0) = \exp\left[-\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} dt'\,H(t')\right], \tag{3.15}$$

where use has been made of Equations (3.5) and (3.6). (Here, we assume that Hamiltonian operators evaluated at different times commute with one another.) The fact that H is undetermined to an arbitrary real additive constant leaves T undetermined to a phase-factor. Incidentally, in the above analysis, time is *not* an operator (we cannot observe time, as such), it is just a parameter (or, more accurately, a continuous label).

3.2 Heisenberg Equation of Motion

We have seen that in the Schrödinger scheme the dynamical variables of the system remain fixed during a period of undisturbed motion, whereas the state kets evolve according to Equation (3.10). However, this is not the only way in which to represent the time evolution of the system.

Suppose that a general state ket A is subject to the transformation

$$|A_t\rangle = T^\dagger(t, t_0)\,|A\rangle. \tag{3.16}$$

This is a time-dependent transformation, because the operator $T(t, t_0)$ obviously depends on time. The subscript t is used to remind us that the transformation is time-dependent. The time evolution of the transformed state ket is given by

$$|A_t t\rangle = T^\dagger(t, t_0)\,|At\rangle = T^\dagger(t, t_0)\,T(t, t_0)\,|At_0\rangle = |A_t t_0\rangle, \tag{3.17}$$

where use has been made of Equations (3.3), (3.5), and the fact that $T(t_0, t_0) = 1$. Clearly, the transformed state ket *does not* evolve in time. Thus, the transformation (3.16) has the effect of bringing all kets representing states of undisturbed motion of the system to rest.

The transformation must also be applied to bras. The dual of Equation (3.16) yields

$$\langle A_t| = \langle A|\,T. \tag{3.18}$$

The transformation rule for a general observable v is obtained from the requirement that the expectation value $\langle A|\,v\,|A\rangle$ should remain invariant. It is easily seen that

$$v_t = T^\dagger\,v\,T. \tag{3.19}$$

Thus, a dynamical variable, which corresponds to a fixed linear operator in Schrödinger's scheme, corresponds to a moving linear operator in this new scheme. It is clear that the transformation (3.16) leads us to a scenario in which the state of the system is represented by a fixed vector, and the dynamical variables are represented by moving linear operators. This is termed the *Heisenberg picture*, as opposed to the *Schrödinger picture*, which is outlined in Section 3.1.

Consider a dynamical variable v corresponding to a fixed linear operator in the Schrödinger picture. According to Equation (3.19), we can write

$$T\,v_t = v\,T. \tag{3.20}$$

Differentiation with respect to time yields

$$\frac{dT}{dt}\,v_t + T\,\frac{dv_t}{dt} = v\,\frac{dT}{dt}. \tag{3.21}$$

With the help of Equation (3.14), this reduces to

$$H\,T\,v_t + \mathrm{i}\,\hbar\,T\,\frac{dv_t}{dt} = v\,H\,T, \tag{3.22}$$

or

$$\mathrm{i}\,\hbar\,\frac{dv_t}{dt} = T^\dagger\,v\,H\,T - T^\dagger\,H\,T\,v_t = v_t\,H_t - H_t\,v_t, \tag{3.23}$$

where

$$H_t = T^\dagger\,H\,T. \tag{3.24}$$

Equation (3.23) can be written

$$\mathrm{i}\,\hbar\,\frac{dv_t}{dt} = [v_t, H_t]. \tag{3.25}$$

Equation (3.25) shows how the dynamical variables of the system evolve in the Heisenberg picture. It is denoted the *Heisenberg equation of motion*. Note that the time-varying dynamical variables in the Heisenberg picture are usually called *Heisenberg dynamical variables* to distinguish them from *Schrödinger dynamical variables* (i.e., the corresponding variables in the Schrödinger picture), which do not evolve in time.

According to Equation (2.22), the Heisenberg equation of motion can be written

$$\frac{dv_t}{dt} = [v_t, H_t]_{qm}, \tag{3.26}$$

where $[\cdots]_{qm}$ denotes the quantum Poisson bracket. Let us compare this equation with the classical time evolution equation for a general dynamical variable v, which can be written in the form [see Equation (2.7)]

$$\frac{dv}{dt} = [v, H]_{cl}. \tag{3.27}$$

Here, $[\cdots]_{cl}$ is the classical Poisson bracket, and H denotes the classical Hamiltonian. The strong resemblance between Equations (3.26) and (3.27) provides us with further justification for our identification of the linear operator H with the energy of the system in quantum mechanics.

Note that if the Hamiltonian does not explicitly depend on time (i.e., the system is not subject to some time-dependent external force) then Equation (3.15) yields

$$T(t, t_0) = \exp\left[-\mathrm{i}\,H\,(t - t_0)/\hbar\right]. \tag{3.28}$$

This operator manifestly commutes with H, so

$$H_t = T^\dagger\,H\,T = H. \tag{3.29}$$

Furthermore, Equation (3.25) gives

$$\mathrm{i}\,\hbar\,\frac{dH}{dt} = [H, H] = 0. \tag{3.30}$$

Thus, if the energy of the system has no explicit time-dependence then it is represented by the same non-time-varying operator H in both the Schrödinger and Heisenberg pictures.

Suppose that v is an observable that commutes with the Hamiltonian (and, hence, with the time evolution operator T). It follows from Equation (3.19) that $v_t = v$. Heisenberg's equation of motion yields

$$\mathrm{i}\,\hbar\,\frac{dv}{dt} = [v, H] = 0. \tag{3.31}$$

Thus, *any observable that commutes with the Hamiltonian is a constant of the motion* (hence, it is represented by the same fixed operator in both the Schrödinger and Heisenberg pictures). Only those observables that *do not* commute with the Hamiltonian evolve in time in the Heisenberg picture.

3.3 Ehrenfest Theorem

We have now introduced all of the basic elements of quantum mechanics. The only thing which is lacking is some rule to determine the form of the quantum mechanical Hamiltonian. For a physical system that possess a classical analogue, we generally assume that the Hamiltonian has the same form as in classical physics (i.e., we replace the classical coordinates and conjugate momenta by the corresponding quantum mechanical operators). This scheme guarantees that quantum mechanics yields the correct classical equations of motion in the classical limit. Whenever an ambiguity arises because of non-commuting observables, this can usually be resolved by requiring the Hamiltonian H to be an Hermitian operator. For instance, we would write the quantum mechanical analogue of the classical product $x\,p_x$, appearing in the Hamiltonian, as the Hermitian product $(1/2)(x\,p_x + p_x\,x)$. When the system in question has no classical analogue then we are reduced to guessing a form for H that reproduces the observed behavior of the system.

Consider a three-dimensional system characterized by three independent Cartesian position coordinates x_i (where i runs from 1 to 3), with three corresponding conjugate momenta p_i. These are represented by three commuting position operators x_i, and three commuting momentum operators p_i, respectively. The commutation relations satisfied by the position and momentum operators are [see Equation (2.25)]

$$[x_i, p_j] = \mathrm{i}\,\hbar\,\delta_{ij}. \tag{3.32}$$

It is helpful to denote (x_1, x_2, x_3) as $\mathbf{x}$ and (p_1, p_2, p_3) as $\mathbf{p}$. The following useful formulae,

$$[x_i, F(\mathbf{x}, \mathbf{p})] = \mathrm{i}\,\hbar\,\frac{\partial F}{\partial p_i}, \tag{3.33}$$

$$[p_i, G(\mathbf{x}, \mathbf{p})] = -\mathrm{i}\,\hbar\,\frac{\partial G}{\partial x_i}, \tag{3.34}$$

where $F(\mathbf{x}, \mathbf{p})$ and $G(\mathbf{x}, \mathbf{p})$ are functions that can be expanded as power series, are easily proved using the fundamental commutation relations, (3.32).

Let us now consider the three-dimensional motion of a free particle of mass m in the Heisenberg picture. The Hamiltonian is assumed to have the same form as in classical physics: i.e.,

$$H(\mathbf{x}, \mathbf{p}) = \frac{p^2}{2\,m} = \frac{1}{2\,m}\sum_{i=1,3} p_i^{\,2}. \tag{3.35}$$

In the following, all dynamical variables are assumed to be Heisenberg dynamical variables, although we will omit the subscript t for the sake of clarity. The time evolution of the momentum operator p_i follows from the Heisenberg equation of motion, (3.25). We find that

$$\frac{dp_i}{dt} = \frac{1}{\mathrm{i}\,\hbar}\,[p_i, H] = 0, \tag{3.36}$$

since p_i automatically commutes with any function of the momentum operators. Thus, for a free particle, the momentum operators are constants of the motion, which means that $p_i(t) = p_i(0)$ at all times t (for i is 1 to 3). The time evolution of the position operator x_i is given by

$$\frac{dx_i}{dt} = \frac{1}{\mathrm{i}\,\hbar}\,[x_i, H] = \frac{1}{\mathrm{i}\,\hbar}\,\frac{1}{2\,m}\,\mathrm{i}\,\hbar\,\frac{\partial}{\partial p_i}\left(\sum_{j=1,3} p_j^{\,2}\right) = \frac{p_i}{m} = \frac{p_i(0)}{m}, \tag{3.37}$$

where use has been made of Equation (3.33). It follows that

$$x_i(t) = x_i(0) + \left[\frac{p_i(0)}{m}\right] t, \tag{3.38}$$

which is analogous to the equation of motion of a classical free particle. Note that even though

$$[x_i(0), x_j(0)] = 0, \tag{3.39}$$

where the position operators are evaluated at equal times, the x_i *do not* commute when evaluated at different times. For instance,

$$[x_i(t), x_i(0)] = \left[\frac{p_i(0)\,t}{m}, x_i(0)\right] = \frac{-\mathrm{i}\,\hbar\,t}{m}. \tag{3.40}$$

Combining the above commutation relation with the uncertainty relation (1.83) yields

$$\langle(\varDelta x_i)^2\rangle_t \langle(\varDelta x_i)^2\rangle_{t=0} \geq \frac{\hbar^2\,t^2}{4\,m^2}. \tag{3.41}$$

This result implies that even if a particle is well-localized at $t = 0$, its position becomes progressively more uncertain with time. This conclusion can also be obtained by studying the propagation of wavepackets in wave mechanics.

Let us now add a potential $V(\mathbf{x})$ to our free particle Hamiltonian:

$$H(\mathbf{x}, \mathbf{p}) = \frac{p^2}{2\,m} + V(\mathbf{x}). \tag{3.42}$$

Here, V is some (real) function of the x_i operators. The Heisenberg equation of motion gives

$$\frac{dp_i}{dt} = \frac{1}{\mathrm{i}\,\hbar}\,[p_i, V(\mathbf{x})] = -\frac{\partial V(\mathbf{x})}{\partial x_i}, \tag{3.43}$$

where use has been made of Equation (3.34). On the other hand, the result

$$\frac{dx_i}{dt} = \frac{p_i}{m} \tag{3.44}$$

still holds, because the x_i all commute with the new term, $V(\mathbf{x})$, in the Hamiltonian. We can use the Heisenberg equation of motion a second time to deduce that

$$\frac{d^2x_i}{dt^2} = \frac{1}{\mathrm{i}\,\hbar}\left[\frac{dx_i}{dt}, H\right] = \frac{1}{\mathrm{i}\,\hbar}\left[\frac{p_i}{m}, H\right] = \frac{1}{m}\,\frac{dp_i}{dt} = -\frac{1}{m}\,\frac{\partial V(\mathbf{x})}{\partial x_i}. \tag{3.45}$$

In vectorial form, this equation becomes

$$m\,\frac{d^2\mathbf{x}}{dt^2} = \frac{d\mathbf{p}}{dt} = -\nabla V(\mathbf{x}). \tag{3.46}$$

This is the quantum mechanical equivalent of Newton's second law of motion. Taking the expectation values of both sides with respect to a Heisenberg state ket that does *not* evolve in time, we obtain the so-called *Ehrenfest theorem*:

$$m\,\frac{d^2\langle\mathbf{x}\rangle}{dt^2} = \frac{d\langle\mathbf{p}\rangle}{dt} = -\langle\nabla V(\mathbf{x})\rangle. \tag{3.47}$$

When written in terms of expectation values, this result is independent of whether we are using the Heisenberg or Schrödinger picture. By contrast, the operator equation (3.46) only holds if $\mathbf{x}$ and $\mathbf{p}$ are understood to be Heisenberg dynamical variables. Note that Equation (3.47) has no dependence on $\hbar$. In fact, it guarantees to us that the centre of a wavepacket always moves like a classical particle.

3.4 Schrödinger Wave Equation

Consider the motion of a particle in three dimensions in the Schrödinger picture. The fixed dynamical variables of the system are the position operators, $\mathbf{x} \equiv (x_1, x_2, x_3)$, and the momentum operators, $\mathbf{p} \equiv (p_1, p_2, p_3)$. The state of the system is represented as some time evolving ket $|At\rangle$.

Let $|\mathbf{x}'\rangle$ represent a simultaneous eigenket of the position operators belonging to the eigenvalues $\mathbf{x}' \equiv (x_1', x_2', x_3')$. Note that, because the position operators are *fixed* in the Schrödinger picture, we do not expect the $|\mathbf{x}'\rangle$ to evolve in time. The wavefunction of the system at time t is defined

$$\psi(\mathbf{x}', t) = \langle\mathbf{x}'|At\rangle. \tag{3.48}$$

The Hamiltonian of the system is taken to be

$$H(\mathbf{x}, \mathbf{p}) = \frac{p^2}{2\,m} + V(\mathbf{x}). \tag{3.49}$$

The Schrödinger equation of motion, (3.10), yields

$$\mathrm{i}\,\hbar\,\frac{\partial\langle\mathbf{x}'|At\rangle}{\partial t} = \langle\mathbf{x}'|\,H\,|At\rangle, \tag{3.50}$$

where use has been made of the time independence of the $|\mathbf{x}'\rangle$. We adopt the Schrödinger representation in which the momentum conjugate to the position operator x_i is written [see Equation (2.74)]

$$p_i = -\mathrm{i}\,\hbar\,\frac{\partial}{\partial x_i}. \tag{3.51}$$

Thus,

$$\left\langle\mathbf{x}'\left|\frac{p^2}{2\,m}\right|At\right\rangle = -\left(\frac{\hbar^2}{2\,m}\right)\nabla'^{\,2}\langle\mathbf{x}'|At\rangle, \tag{3.52}$$

where use has been made of Equation (2.78). Here, $\nabla' \equiv (\partial/\partial x', \partial/\partial y', \partial/\partial z')$ denotes the gradient operator written in terms of the position eigenvalues. We can also write

$$\langle \mathbf{x}'|\, V(\mathbf{x}) = V(\mathbf{x}')\,\langle \mathbf{x}'|, \tag{3.53}$$

where $V(\mathbf{x}')$ is a scalar function of the position eigenvalues. Combining Equations (3.49), (3.50), (3.52), and (3.53), we obtain

$$\mathrm{i}\,\hbar\,\frac{\partial\langle \mathbf{x}'|At\rangle}{\partial t} = -\left(\frac{\hbar^2}{2\,m}\right)\nabla'^{\,2}\langle \mathbf{x}'|At\rangle + V(\mathbf{x}')\langle \mathbf{x}'|At\rangle, \tag{3.54}$$

which can also be written

$$\mathrm{i}\,\hbar\,\frac{\partial\psi(\mathbf{x}', t)}{\partial t} = -\left(\frac{\hbar^2}{2\,m}\right)\nabla'^{\,2}\psi(\mathbf{x}', t) + V(\mathbf{x}')\,\psi(\mathbf{x}', t). \tag{3.55}$$

This is Schrödinger's famous wave equation, and is the basis of wave mechanics. Note, however, that the wave equation is just one of many possible representations of quantum mechanics. It just happens to give a type of equation that we know how to solve. In deriving the wave equation, we have chosen to represent the system in terms of the eigenkets of the position operators, instead of those of the momentum operators. We have also fixed the relative phases of the $|\mathbf{x}'\rangle$ according to the Schrödinger representation, so that Equation (3.51) is valid. Finally, we have chosen to work in the Schrödinger picture, in which state kets evolve and dynamical variables are fixed, instead of the Heisenberg picture, in which the opposite is true.

Suppose that the ket $|At\rangle$ is an eigenket of the Hamiltonian belonging to the eigenvalue H': i.e.,

$$H\,|At\rangle = H'\,|At\rangle. \tag{3.56}$$

The Schrödinger equation of motion, (3.10), yields

$$\mathrm{i}\,\hbar\,\frac{d|At\rangle}{dt} = H'\,|At\rangle. \tag{3.57}$$

This can be integrated to give

$$|At\rangle = \exp[-\mathrm{i}\,H'(t - t_0)/\hbar]\,|At_0\rangle. \tag{3.58}$$

Note that $|At\rangle$ only differs from $|At_0\rangle$ by a phase-factor. The direction of the vector remains fixed in ket space. This suggests that if the system is initially in an eigenstate of the Hamiltonian then it remains in this state for ever, as long as the system is undisturbed. Such a state is called a *stationary state*. The wavefunction of a stationary state satisfies

$$\psi(\mathbf{x}', t) = \psi(\mathbf{x}', t_0)\exp[-\mathrm{i}\,H'\,(t - t_0)/\hbar]. \tag{3.59}$$

Substituting the above relation into the Schrödinger wave equation, (3.55), we obtain

$$-\left(\frac{\hbar^2}{2\,m}\right)\nabla'^{\,2}\psi_0(\mathbf{x}') + [V(\mathbf{x}') - E]\,\psi_0(\mathbf{x}') = 0, \tag{3.60}$$

where $\psi_0(\mathbf{x}') \equiv \psi(\mathbf{x}', t_0)$, and $E = H'$ is the energy of the system. This is Schrödinger's time-independent wave equation. A *bound state* solution of the above equation, in which the particle is confined within a finite region of space, satisfies the boundary condition

$$\psi_0(\mathbf{x}') \to 0 \qquad \text{as } |\mathbf{x}'| \to \infty. \tag{3.61}$$

Such a solution is only possible if

$$E < \lim_{|\mathbf{x}'|\to\infty} V(\mathbf{x}'). \tag{3.62}$$

Since it is conventional to set the potential at infinity equal to zero, the above relation implies that bound states are equivalent to negative energy states. The boundary condition (3.61) is sufficient to uniquely specify the solution of Equation (3.60).

The quantity $\rho(\mathbf{x}', t)$, defined by

$$\rho(\mathbf{x}', t) = |\psi(\mathbf{x}', t)|^2, \tag{3.63}$$

is termed the *probability density*. Recall, from Equation (2.30), that the probability of observing the particle in some volume element d^3x' around position $\mathbf{x}'$ is proportional to $\rho(\mathbf{x}', t)\, d^3x'$. The probability is *equal* to $\rho(\mathbf{x}', t)\, d^3x'$ if the wavefunction is properly normalized, so that

$$\int d^3x'\, \rho(\mathbf{x}', t) = 1. \tag{3.64}$$

The Schrödinger time-dependent wave equation, (3.55), can easily be transformed into a conservation equation for the probability density:

$$\frac{\partial \rho}{\partial t} + \nabla' \cdot \mathbf{j} = 0. \tag{3.65}$$

The *probability current*, $\mathbf{j}$, takes the form

$$\mathbf{j}(\mathbf{x}', t) = -\left(\frac{\mathrm{i}\,\hbar}{2\,m}\right)[\psi^*\,\nabla'\psi - (\nabla'\psi^*)\,\psi] = \left(\frac{\hbar}{m}\right)\mathrm{Im}(\psi^*\,\nabla'\psi). \tag{3.66}$$

We can integrate Equation (3.65) over all space, using the divergence theorem, and the boundary condition $\rho \to 0$ as $|\mathbf{x}'| \to \infty$, to obtain

$$\frac{d}{dt}\int d^3x'\, \rho(\mathbf{x}', t) = 0. \tag{3.67}$$

Thus, the Schrödinger wave equation *conserves* probability. In particular, if the wavefunction starts off properly normalized, according to Equation (3.64), then it remains properly normalized at all subsequent times. It is easily demonstrated that

$$\int d^3x'\, \mathbf{j}(\mathbf{x}', t) = \frac{\langle \mathbf{p} \rangle_t}{m}, \tag{3.68}$$

where $\langle \mathbf{p} \rangle_t$ denotes the expectation value of the momentum evaluated at time t. Clearly, the probability current is indirectly related to the particle momentum.

In deriving Equations (3.65), we have, naturally, assumed that the potential $V(\mathbf{x}')$ is real. Suppose, however, that the potential has an imaginary component. In this case, Equation (3.65) generalizes to

$$\frac{\partial\rho}{\partial t} + \nabla'\cdot\mathbf{j} = \frac{2\,\mathrm{Im}(V)}{\hbar}\,\rho, \tag{3.69}$$

giving

$$\frac{\partial}{\partial t}\int d^3x'\,\rho(\mathbf{x}',t) = \frac{2}{\hbar}\,\mathrm{Im}\int d^3x'\,V(\mathbf{x}')\,\rho(\mathbf{x}',t). \tag{3.70}$$

Thus, if $\mathrm{Im}(V) < 0$ then the total probability of observing the particle anywhere in space decreases monotonically with time. Thus, an imaginary potential can be used to account for the disappearance of a particle. Such a potential is often employed to model nuclear reactions in which incident particles are absorbed by nuclei.

Exercises

3.1 Let $\mathbf{x} \equiv (x_1, x_2, x_3)$ be a set of Cartesian position operators, and let $\mathbf{p} \equiv (p_1, p_2, p_3)$ be the corresponding momentum operators. Demonstrate that

$$[x_i, F(\mathbf{x},\mathbf{p})] = \mathrm{i}\,\hbar\,\frac{\partial F}{\partial p_i},$$
$$[p_i, G(\mathbf{x},\mathbf{p})] = -\mathrm{i}\,\hbar\,\frac{\partial G}{\partial x_i},$$

where $i = 1, 3$, and $F(\mathbf{x},\mathbf{p})$, $G(\mathbf{x},\mathbf{p})$ are functions that can be expanded as power series.

3.2 Assuming that the potential $V(\mathbf{x})$ is complex, demonstrate that the Schrödinger time-dependent wave equation, (3.55), can be transformed to give

$$\frac{\partial\rho}{\partial t} + \nabla'\cdot\mathbf{j} = 2\,\frac{\mathrm{Im}(V)}{\hbar}\,\rho,$$

where

$$\rho(\mathbf{x}',t) = |\psi(\mathbf{x}',t)|^2,$$

and

$$\mathbf{j}(\mathbf{x}',t) = \left(\frac{\hbar}{m}\right)\mathrm{Im}(\psi^*\,\nabla'\psi).$$

3.3 Consider one-dimensional quantum harmonic oscillator whose Hamiltonian is

$$H = \frac{p_x^2}{2\,m} + \frac{1}{2}\,m\,\omega^2\,x^2,$$

where x and p_x are conjugate position and momentum operators, respectively, and m, ω are positive constants.

(a) Demonstrate that the expectation value of H, for a general state, is positive definite.

(b) Let

$$A = \sqrt{\frac{m\,\omega}{2\,\hbar}}\,x + {\rm i}\,\frac{p_x}{\sqrt{2\,m\,\omega\,\hbar}}.$$

Deduce that

$$[A, A^\dagger] = 1,$$

$$H = \hbar\,\omega\left(\frac{1}{2} + A^\dagger\,A\right),$$

$$[H, A] = -\hbar\,\omega\,A,$$

$$[H, A^\dagger] = \hbar\,\omega\,A^\dagger.$$

(c) Suppose that $|E\rangle$ is an eigenket of the Hamiltonian whose corresponding energy is E: i.e.,

$$H\,|E\rangle = E\,|E\rangle.$$

Demonstrate that

$$H\,A\,|E\rangle = (E - \hbar\,\omega)\,A\,|E\rangle,$$

$$H\,A^\dagger\,|E\rangle = (E + \hbar\,\omega)\,A^\dagger\,|E\rangle.$$

Hence, deduce that the allowed values of E are

$$E_n = (n + 1/2)\,\hbar\,\omega,$$

where $n = 0, 1, 2, \cdots$.

(d) Let $|E_n\rangle$ be a properly normalized (i.e., $\langle E_n|E_n\rangle = 1$) energy eigenket corresponding to the eigenvalue E_n. Show that the kets can be defined such that

$$A\,|E_n\rangle = \sqrt{n}\,|E_{n-1}\rangle,$$

$$A^\dagger\,|E_n\rangle = \sqrt{n+1}\,|E_{n+1}\rangle.$$

Hence, deduce that

$$|E_n\rangle = \frac{1}{\sqrt{n!}}\,(A^\dagger)^n\,|E_0\rangle.$$

(e) Let the $\psi_n(x') = \langle x'|E_n\rangle$ be the wavefunctions of the properly normalized energy eigenkets. Given that

$$A\,|E_0\rangle = |0\rangle,$$

deduce that

$$\left(\frac{x'}{x_0} + x_0\,\frac{d}{dx'}\right)\psi_0(x') = 0,$$

where $x_0 = (\hbar/m\,\omega)^{1/2}$. Hence, show that

$$\psi_n(x') = \frac{1}{\pi^{1/4}\,(2^n\,n!)^{1/2}\,x_0^{\,n+1/2}}\left(x' - x_0^2\,\frac{d}{dx'}\right)^n \exp\left[-\frac{1}{2}\left(\frac{x'}{x_0}\right)^2\right].$$

3.4 Consider the one-dimensional quantum harmonic oscillator discussed in Exercise 3.3. Let $|n\rangle$ be a properly normalized energy eigenket belonging to the eigenvalue E_n. Show that

$$\langle n'|\,x\,|n\rangle = \sqrt{\frac{\hbar}{2\,m\,\omega}}\left(\sqrt{n}\,\delta_{n'\,n-1} + \sqrt{n+1}\,\delta_{n'\,n+1}\right),$$

$$\langle n'|\,p_x\,|n\rangle = \mathrm{i}\sqrt{\frac{m\,\hbar\,\omega}{2}}\left(-\sqrt{n}\,\delta_{n'\,n-1} + \sqrt{n+1}\,\delta_{n'\,n+1}\right),$$

$$\langle n'|\,x^2\,|n\rangle = \left(\frac{\hbar}{2\,m\,\omega}\right)\left(\sqrt{n\,(n-1)}\,\delta_{n'\,n-2} + \sqrt{(n+1)\,(n+2)}\,\delta_{n'\,n+2} + (2\,n+1)\,\delta_{n'\,n}\right),$$

$$\langle n'|\,p_x^2\,|n\rangle = \left(\frac{m\,\hbar\,\omega}{2}\right)\left(-\sqrt{n\,(n-1)}\,\delta_{n'\,n-2} - \sqrt{(n+1)\,(n+2)}\,\delta_{n'\,n+2} + (2\,n+1)\,\delta_{n',n}\right).$$

Hence, deduce that

$$\langle(\Delta x)^2\rangle\,\langle(\Delta p_x)^2\rangle = (n+1/2)^2\,\hbar^2$$

for the *n*th eigenstate.

3.5 Consider a particle in one dimension whose Hamiltonian is

$$H = \frac{p_x^2}{2\,m} + V(x).$$

By calculating $[[H, x], x]$, demonstrate that

$$\sum_{n'} |\langle n|\,x\,|n'\rangle|^2\,(E_{n'} - E_n) = \frac{\hbar^2}{2\,m},$$

where $|n\rangle$ is a properly normalized energy eigenket corresponding to the eigenvalue E_n, and the sum is over all eigenkets.

3.6 Consider a particle in one dimension whose Hamiltonian is

$$H = \frac{p_x^2}{2\,m} + V(x).$$

Suppose that the potential is periodic, such that

$$V(x-a) = V(x),$$

for all x. Deduce that

$$[D(a), H] = 0,$$

where $D(a)$ is the displacement operator defined in Exercise 2.4. Hence, show that the wavefunction of an energy eigenstate has the general form

$$\psi(x') = \mathrm{e}^{\,\mathrm{i}\,k'\,x'}\,u(x'),$$

where k' is a real parameter, and $u(x'-a) = u(x')$ for all x'. This result is known as the *Bloch theorem*.

3.7 Consider the one-dimensional quantum harmonic oscillator discussed in Exercise 3.3. Show that the Heisenberg equations of motion of the ladder operators A and $A^\dagger$ are

$$\frac{dA}{dt} = -\mathrm{i}\,\omega\,A,$$
$$\frac{dA^\dagger}{dt} = \mathrm{i}\,\omega\,A^\dagger,$$

respectively. Hence, deduce that the momentum and position operators evolve in time as

$$p_x(t) = \cos(\omega\,t)\,p_x(0) - m\,\omega\,\sin(\omega\,t)\,x(0),$$
$$x(t) = \cos(\omega\,t)\,x(0) + \frac{\sin(\omega\,t)}{m\,\omega}\,p_x(0),$$

respectively, in the Heisenberg picture.

3.8 Consider a one-dimensional stationary bound state. Using the time-independent Schrödinger equation, prove that

$$\left\langle\frac{p_x^{\,2}}{2\,m}\right\rangle = E - \langle V\rangle,$$

and

$$\left\langle\frac{p_x^{\,2}}{2\,m}\right\rangle = -E + \langle V\rangle + \left\langle x\,\frac{dV}{dx}\right\rangle.$$

[Hint: You can assume, without loss of generality, that the stationary wavefunction is real.] Hence, prove the *Virial theorem*,

$$\left\langle\frac{p_x^{\,2}}{2\,m}\right\rangle = \frac{1}{2}\left\langle x\,\frac{dV}{dx}\right\rangle.$$

4 Orbital Angular Momentum

4.1 Orbital Angular Momentum

Consider a particle described by the Cartesian coordinates $(x, y, z) \equiv \mathbf{x}$ and their conjugate momenta $(p_x, p_y, p_z) \equiv \mathbf{p}$. The classical definition of the orbital angular momentum of such a particle about the origin is $\mathbf{L} = \mathbf{x} \times \mathbf{p}$, giving

$$L_x = y\,p_z - z\,p_y, \tag{4.1}$$

$$L_y = z\,p_x - x\,p_z, \tag{4.2}$$

$$L_z = x\,p_y - y\,p_x. \tag{4.3}$$

Let us assume that the operators $(L_x, L_y, L_z) \equiv \mathbf{L}$ which represent the components of orbital angular momentum in quantum mechanics can be defined in an analogous manner to the corresponding components of classical angular momentum. In other words, we are going to assume that the above equations specify the angular momentum operators in terms of the position and linear momentum operators. Note that L_x, L_y, and L_z are Hermitian, so they represent things which can, in principle, be measured. Note, also, that there is no ambiguity regarding the order in which operators appear in products on the right-hand sides of Equations (4.1)–(4.3), because all of the products consist of operators that commute.

The fundamental commutation relations satisfied by the position and linear momentum operators are [see Equations (2.23)–(2.25)]

$$[x_i, x_j] = 0, \tag{4.4}$$

$$[p_i, p_j] = 0, \tag{4.5}$$

$$[x_i, p_j] = \mathrm{i}\,\hbar\,\delta_{ij}, \tag{4.6}$$

where i and j stand for either x, y, or z. Consider the commutator of the operators L_x and L_z:

$$\begin{aligned}[L_x, L_y] &= [(y\,p_z - z\,p_y), (z\,p_x - x\,p_z)] = y\,[p_z, z]\,p_x + x\,p_y\,[z, p_z]\\ &= \mathrm{i}\,\hbar\,(-y\,p_x + x\,p_y) = \mathrm{i}\,\hbar\,L_z.\end{aligned} \tag{4.7}$$

The cyclic permutations of the above result yield the fundamental commutation relations satisfied by the components of an orbital angular momentum:

$$[L_x, L_y] = \mathrm{i}\,\hbar\,L_z, \tag{4.8}$$

$$[L_y, L_z] = \mathrm{i}\,\hbar\,L_x, \tag{4.9}$$

$$[L_z, L_x] = \mathrm{i}\,\hbar\,L_y. \tag{4.10}$$

These can be summed up more succinctly by writing

$$\mathbf{L} \times \mathbf{L} = \mathrm{i}\,\hbar\,\mathbf{L}. \tag{4.11}$$

The three commutation relations (4.8)–(4.10) are the foundation for the whole theory of angular momentum in quantum mechanics. Whenever we encounter three operators having these commutation relations, we know that the dynamical variables that they represent have identical properties to those of the components of an angular momentum (which we are about to derive). In fact, we shall assume that *any three operators that satisfy the commutation relations (4.8)–(4.10) represent the components of some sort of angular momentum.*

Suppose that there are N particles in the system, with angular momentum vectors $\mathbf{L}_i$ (where i runs from 1 to N). Each of these vectors satisfies Equation (4.11), so that

$$\mathbf{L}_i \times \mathbf{L}_i = \mathrm{i}\,\hbar\,\mathbf{L}_i. \tag{4.12}$$

However, we expect the angular momentum operators belonging to different particles to commute, because they represent different degrees of freedom of the system. So, we can write

$$\mathbf{L}_i \times \mathbf{L}_j + \mathbf{L}_j \times \mathbf{L}_i = 0, \tag{4.13}$$

for $i \neq j$. Consider the total angular momentum of the system, $\mathbf{L} = \sum_{i=1,N} \mathbf{L}_i$. It is clear from Equations (4.12) and (4.13) that

$$\begin{aligned}\mathbf{L} \times \mathbf{L} &= \sum_{i=1,N} \mathbf{L}_i \times \sum_{j=1,N} \mathbf{L}_j = \sum_{i=1,N} \mathbf{L}_i \times \mathbf{L}_i + \frac{1}{2} \sum_{i,j=1,N}^{i\neq j} (\mathbf{L}_i \times \mathbf{L}_j + \mathbf{L}_j \times \mathbf{L}_i)\\ &= \mathrm{i}\,\hbar \sum_{i=1,N} \mathbf{L}_i = \mathrm{i}\,\hbar\,\mathbf{L}.\end{aligned} \tag{4.14}$$

Thus, the sum of two or more angular momentum vectors satisfies the same commutation relation as a primitive angular momentum vector. In particular, the total angular momentum of the system satisfies the commutation relation (4.11).

The immediate conclusion which can be drawn from the commutation relations (4.8)–(4.10) is that the three components of an angular momentum vector cannot be specified (or measured) simultaneously. In fact, once we have specified one component, the values of other two components become uncertain. It is conventional to specify the z-component, L_z.

Consider the magnitude squared of the angular momentum vector, $L^2 \equiv L_x^2 + L_y^2 + L_z^2$. The commutator of L^2 and L_z is written

$$[L^2, L_z] = [L_x^2, L_z] + [L_y^2, L_z] + [L_z^2, L_z]. \tag{4.15}$$

It is easily demonstrated that

$$[L_x^2, L_z] = -\mathrm{i}\,\hbar\,(L_x\,L_y + L_y\,L_x), \tag{4.16}$$

$$[L_y^2, L_z] = +\mathrm{i}\,\hbar\,(L_x\,L_y + L_y\,L_x), \tag{4.17}$$

$$[L_z^2, L_z] = 0, \tag{4.18}$$

so

$$[L^2, L_z] = 0. \tag{4.19}$$

Because there is nothing special about the z-direction, we conclude that L^2 also commutes with L_x and L_y. It is clear from Equations (4.8)–(4.10) and (4.19) that the best we can do in quantum mechanics is to specify the magnitude of an angular momentum vector along with *one* of its components (by convention, the z-component).

It is convenient to define the *shift operators* L^+ and L^-:

$$L^+ = L_x + \mathrm{i}\,L_y, \tag{4.20}$$

$$L^- = L_x - \mathrm{i}\,L_y. \tag{4.21}$$

It can easily be shown that

$$[L^+, L_z] = -\hbar\,L^+, \tag{4.22}$$

$$[L^-, L_z] = +\hbar\,L^-, \tag{4.23}$$

$$[L^+, L^-] = 2\,\hbar\,L_z, \tag{4.24}$$

and also that both shift operators commute with L^2.

4.2 Eigenvalues of Orbital Angular Momentum

Suppose that the simultaneous eigenkets of L^2 and L_z are completely specified by two quantum numbers, l and m. These kets are denoted $|l, m\rangle$. The quantum number m is defined by

$$L_z\,|l, m\rangle = m\,\hbar\,|l, m\rangle. \tag{4.25}$$

Thus, m is the eigenvalue of L_z divided by $\hbar$. It is possible to write such an equation because $\hbar$ has the dimensions of angular momentum. Note that m is a real number, because L_z is an Hermitian operator.

We can write

$$L^2\,|l, m\rangle = f(l, m)\,\hbar^2\,|l, m\rangle, \tag{4.26}$$

without loss of generality, where $f(l, m)$ is some real dimensionless function of l and m. Later on, we will show that $f(l, m) = l\,(l + 1)$. Now,

$$\langle l, m|\,L^2 - L_z^2\,|l, m\rangle = \langle l, m|\,f(l, m)\,\hbar^2 - m^2\,\hbar^2\,|l, m\rangle = [f(l, m) - m^2]\,\hbar^2, \tag{4.27}$$

assuming that the $|l, m\rangle$ have unit norms. However,

$$\langle l, m|\,L^2 - L_z^2\,|l, m\rangle = \langle l, m|\,L_x^2 + L_y^2\,|l, m\rangle = \langle l, m|\,L_x^2\,|l, m\rangle + \langle l, m|\,L_y^2\,|l, m\rangle. \tag{4.28}$$

It is easily demonstrated that

$$\langle A|\,\xi^2\,|A\rangle \geq 0, \tag{4.29}$$

where $|A\rangle$ is a general ket, and ξ is an Hermitian operator. The proof follows from the observation that

$$\langle A|\,\xi^2\,|A\rangle = \langle A|\,\xi^\dagger\,\xi\,|A\rangle = \langle B|B\rangle, \tag{4.30}$$

where $|B\rangle = \xi\,|A\rangle$, plus the fact that $\langle B|B\rangle \geq 0$ for a general ket $|B\rangle$ [see Equation (1.21)]. It follows from Equations (4.27)–(4.29) that

$$m^2 \leq f(l,m). \tag{4.31}$$

Consider the effect of the shift operator L^+ on the eigenket $|l,m\rangle$. It is easily demonstrated that

$$L^2(L^+|l,m\rangle) = \hbar^2\, f(l,m)\,(L^+|l,m\rangle), \tag{4.32}$$

where use has been made of Equation (4.26), plus the fact that L^2 and L^+ commute. It follows that the ket $L^+|l,m\rangle$ has the same eigenvalue of L^2 as the ket $|l,m\rangle$. Thus, the shift operator L^+ does not affect the magnitude of the angular momentum of any eigenket it acts upon. However,

$$L_z\,L^+|l,m\rangle = (L^+L_z + [L_z,L^+])\,|l,m\rangle = (L^+L_z + \hbar\,L^+)\,|l,m\rangle = (m+1)\,\hbar\,L^+|l,m\rangle, \tag{4.33}$$

where use has been made of Equation (4.22). The above equation implies that $L^+|l,m\rangle$ is proportional to $|l,m+1\rangle$. We can write

$$L^+|l,m\rangle = c_{l\,m}^{+}\,\hbar\,|l,m+1\rangle, \tag{4.34}$$

where $c_{l\,m}^{+}$ is a number. It is clear that if the operator L^+ acts on a simultaneous eigenstate of L^2 and L_z then the eigenvalue of L^2 remains unchanged, but the eigenvalue of L_z is increased by $\hbar$. For this reason, L^+ is called a *raising operator*.

Using similar arguments to those given above, it is possible to demonstrate that

$$L^-\,|l,m\rangle = c_{l\,m}^{-}\,\hbar\,|l,m-1\rangle. \tag{4.35}$$

Hence, L^- is called a *lowering operator*.

The shift operators, L^+ and L^-, respectively step the value of m up and down by unity each time they operate on one of the simultaneous eigenkets of L^2 and L_z. It would appear, at first sight, that any value of m can be obtained by applying these operators a sufficient number of times. However, according to Equation (4.31), there is a definite upper bound to the values that m^2 can take. This bound is determined by the eigenvalue of L^2 [see Equation (4.26)]. It follows that there is a maximum and a minimum possible value which m can take. Suppose that we attempt to raise the value of m above its maximum value $m_{\rm max}$. Since there is no state with $m > m_{\rm max}$, we must have

$$L^+|l,m_{\rm max}\rangle = |0\rangle. \tag{4.36}$$

This implies that

$$L^-\,L^+|l,m_{\rm max}\rangle = |0\rangle. \tag{4.37}$$

However,

$$L^-\,L^+ = L_x^{\,2} + L_y^{\,2} + {\rm i}\,[L_x,L_y] = L^2 - L_z^{\,2} - \hbar\,L_z, \tag{4.38}$$

so Equation (4.37) yields

$$(L^2 - L_z^{\,2} - \hbar\,L_z)\,|l,m_{\rm max}\rangle = |0\rangle. \tag{4.39}$$

The above equation can be rearranged to give

$$L^2\,|l,m_{\rm max}\rangle = (L_z^{\,2} + \hbar\,L_z)\,|l,m_{\rm max}\rangle = m_{\rm max}(m_{\rm max}+1)\,\hbar^2\,|l,m_{\rm max}\rangle. \tag{4.40}$$

Comparison of this equation with Equation (4.26) yields the result

$$f(l, m_{\max}) = m_{\max}(m_{\max} + 1). \tag{4.41}$$

But, when L^- operates on $|n, m_{\max}\rangle$ it generates $|n, m_{\max} - 1\rangle$, $|n, m_{\max} - 2\rangle$, etc. Since the lowering operator does not change the eigenvalue of L^2, all of these states must correspond to the same value of f, namely $m_{\max}(m_{\max} + 1)$. Thus,

$$L^2\,|l, m\rangle = m_{\max}(m_{\max} + 1)\,\hbar^2\,|l, m\rangle. \tag{4.42}$$

At this stage, we can give the unknown quantum number l the value $m_{\max}$, without loss of generality. We can also write the above equation in the form

$$L^2\,|l, m\rangle = l\,(l + 1)\,\hbar^2\,|l, m\rangle. \tag{4.43}$$

It is easily seen that

$$L^-\,L^+\,|l, m\rangle = (L^2 - L_z^2 - \hbar\,L_z)\,|l, m\rangle = \hbar^2\,[l\,(l + 1) - m\,(m + 1)]\,|l, m\rangle. \tag{4.44}$$

Thus,

$$\langle l, m|\,L^-\,L^+\,|l, m\rangle = \hbar^2\,[l\,(l + 1) - m\,(m + 1)]. \tag{4.45}$$

However, we also know that

$$\langle l, m|L^-\,L^+|l, m\rangle = \langle l, m|L^-\,\hbar\,c_{l\,m}^+|l, m + 1\rangle = \hbar^2\,c_{l\,m}^+\,c_{l\,m+1}^-, \tag{4.46}$$

where use has been made of Equations (4.34) and (4.35). It follows that

$$c_{l\,m}^+\,c_{l\,m+1}^- = [l\,(l + 1) - m\,(m + 1)]. \tag{4.47}$$

Consider the following:

$$\begin{aligned}
\langle l, m|\,L^-\,|l, m + 1\rangle &= \langle l, m|\,L_x\,|l, m + 1\rangle - \mathrm{i}\,\langle l, m|\,L_y\,|l, m + 1\rangle \\
&= \langle l, m + 1|\,L_x\,|l, m\rangle^* - \mathrm{i}\,\langle l, m + 1|\,L_y\,|l, m\rangle^* \\
&= (\langle l, m + 1|\,L_x\,|l, m\rangle + \mathrm{i}\,\langle l, m + 1|\,L_y\,|l, m\rangle)^* \\
&= \langle l, m + 1|\,L^+\,|l, m\rangle^*,
\end{aligned} \tag{4.48}$$

where use has been made of the fact that L_x and L_y are Hermitian. The above equation reduces to

$$c_{l\,m+1}^- = (c_{l\,m}^+)^* \tag{4.49}$$

with the aid of Equations (4.34) and (4.35).

Equations (4.47) and (4.49) can be combined to give

$$|c_{l\,m}^+|^2 = [l\,(l + 1) - m\,(m + 1)]. \tag{4.50}$$

The solution of the above equation is

$$c_{l\,m}^{+} = \sqrt{l\,(l+1) - m\,(m+1)}. \tag{4.51}$$

Note that $c_{l\,m}^{+}$ is undetermined to an arbitrary phase-factor [i.e., we can replace $c_{l\,m}^{+}$, given above, by $c_{l\,m}^{+}\exp(\mathrm{i}\,\gamma)$, where γ is real, and we still satisfy Equation (4.50)]. We have made the arbitrary, but convenient, choice that $c_{l\,m}^{+}$ is real and positive. This is equivalent to choosing the relative phases of the eigenkets $|l, m\rangle$. According to Equation (4.49),

$$c_{l\,m}^{-} = (c_{l\,m-1}^{+})^{*} = \sqrt{l\,(l+1) - m\,(m-1)}. \tag{4.52}$$

We have already seen that the inequality (4.31) implies that there is a maximum and a minimum possible value of m. The maximum value of m is denoted l. What is the minimum value? Suppose that we try to lower the value of m below its minimum value $m_{\rm min}$. Because there is no state with $m < m_{\rm min}$, we must have

$$L^{-}\,|l, m_{\rm min}\rangle = 0. \tag{4.53}$$

According to Equation (4.35), this implies that

$$c_{l\,m_{\rm min}}^{-} = 0. \tag{4.54}$$

It can be seen from Equation (4.52) that $m_{\rm min} = -l$. We conclude that m can take a “ladder” of discrete values, each rung differing from its immediate neighbors by unity. The top rung is l, and the bottom rung is $-l$. There are only two possible choices for l. Either it is an integer (e.g., $l = 2$, which allows m to take the values $-2, -1, 0, 1, 2$), or it is a half-integer (e.g., $l = 3/2$, which allows m to take the values $-3/2, -1/2, 1/2, 3/2$). We shall prove in the next section that an orbital angular momentum can only take integer values of l.

In summary, using just the fundamental commutation relations (4.8)–(4.10), plus the fact that L_x, L_y, and L_z are Hermitian operators, we have shown that the eigenvalues of $L^2 \equiv L_x^2 + L_y^2 + L_z^2$ can be written $l\,(l+1)\,\hbar^2$, where l is an integer, or a half-integer. We have also demonstrated that the eigenvalues of L_z can only take the values $m\,\hbar$, where m lies in the range $-l, -l+1, \cdots l-1, l$. Let $|l, m\rangle$ denote a properly normalized simultaneous eigenket of L^2 and L_z, belonging to the eigenvalues $l\,(l+1)\,\hbar^2$ and $m\,\hbar$, respectively. We have shown that

$$L^{+}\,|l, m\rangle = \sqrt{l\,(l+1) - m\,(m+1)}\,\hbar\,|l, m+1\rangle \tag{4.55}$$

$$L^{-}\,|l, m\rangle = \sqrt{l\,(l+1) - m\,(m-1)}\,\hbar\,|l, m-1\rangle, \tag{4.56}$$

where $L^{\pm} = L_x \pm \mathrm{i}\,L_y$ are the so-called shift operators.

4.3 Rotation Operators

Consider a particle whose position is described by the spherical polar coordinates (r, θ, φ). The classical momentum conjugate to the azimuthal angle φ is the z-component of angular momentum,

L_z. According to Section 2.5, in quantum mechanics we can always adopt the Schrödinger representation, for which ket space is spanned by the simultaneous eigenkets of the position operators r, θ, and φ, and L_z takes the form

$$L_z = -\mathrm{i}\,\hbar\,\frac{\partial}{\partial\varphi}. \tag{4.57}$$

We can do this because there is nothing in Section 2.5 which specifies that we have to use Cartesian coordinates—the representation (2.74) works for any well-defined set of coordinates.

Consider an operator $R(\Delta\varphi)$ that rotates the system through an angle $\Delta\varphi$ about the z-axis. This operator is very similar to the operator $D(\Delta x)$, introduced in Section 2.8, which translates the system a distance Δx along the x-axis. We were able to demonstrate in Section 2.8 that

$$p_x = \mathrm{i}\,\hbar\,\lim_{\delta x\to 0}\frac{D(\delta x)-1}{\delta x}, \tag{4.58}$$

where p_x is the linear momentum conjugate to x. There is nothing in our derivation of this result which specifies that x has to be a Cartesian coordinate. Thus, the result should apply just as well to an angular coordinate. We conclude that

$$L_z = \mathrm{i}\,\hbar\,\lim_{\delta\varphi\to 0}\frac{R(\delta\varphi)-1}{\delta\varphi}. \tag{4.59}$$

According to Equation (4.59), we can write

$$R(\delta\varphi) = 1 - \mathrm{i}\,L_z\,\delta\varphi/\hbar \tag{4.60}$$

in the limit $\delta\varphi \to 0$. In other words, the angular momentum operator L_z can be used to rotate the system about the z-axis by an infinitesimal amount. We say that L_z is the *generator* of rotations about the z-axis. The above equation implies that

$$R(\Delta\varphi) = \lim_{N\to\infty}\left(1 - \mathrm{i}\,\frac{\Delta\varphi}{N}\frac{L_z}{\hbar}\right)^N, \tag{4.61}$$

which reduces to

$$R(\Delta\varphi) = \exp(-\mathrm{i}\,L_z\,\Delta\varphi/\hbar). \tag{4.62}$$

Note that $R(\Delta\varphi)$ has all of the properties we would expect of a rotation operator: i.e.,

$$R(0) = 1, \tag{4.63}$$

$$R(\Delta\varphi)\,R(-\Delta\varphi) = 1, \tag{4.64}$$

$$R(\Delta\varphi_1)\,R(\Delta\varphi_2) = R(\Delta\varphi_1 + \Delta\varphi_2). \tag{4.65}$$

Suppose that the system is in a simultaneous eigenstate of L^2 and L_z. As before, this state is represented by the eigenket $|l, m\rangle$, where the eigenvalue of L^2 is $l\,(l+1)\,\hbar^2$, and the eigenvalue of L_z is $m\,\hbar$. We expect the wavefunction to remain unaltered if we rotate the system 2π degrees about the z-axis. Thus,

$$R(2\pi)\,|l,m\rangle = \exp(-\mathrm{i}\,L_z\,2\pi/\hbar)\,|l,m\rangle = \exp(-\mathrm{i}\,2\pi\,m)\,|l,m\rangle = |l,m\rangle. \tag{4.66}$$

We conclude that m must be an integer. This implies, from the previous section, that l must also be an integer. Thus, an *orbital* angular momentum can only take *integer* values of the quantum numbers l and m.

Consider the action of the rotation operator $R(\Delta\varphi)$ on an eigenstate possessing zero angular momentum about the z-axis (i.e., an $m = 0$ state). We have

$$R(\Delta\varphi)\,|l,0\rangle = \exp(0)\,|l,0\rangle = |l,0\rangle. \tag{4.67}$$

Thus, the eigenstate is invariant to rotations about the z-axis. Clearly, its wavefunction must be symmetric about the z-axis.

There is nothing special about the z-axis, so we can write

$$R_x(\Delta\varphi_x) = \exp(-\mathrm{i}\,L_x\,\Delta\varphi_x/\hbar), \tag{4.68}$$

$$R_y(\Delta\varphi_y) = \exp(-\mathrm{i}\,L_y\,\Delta\varphi_y/\hbar), \tag{4.69}$$

$$R_z(\Delta\varphi_y) = \exp(-\mathrm{i}\,L_z\,\Delta\varphi_z/\hbar), \tag{4.70}$$

by analogy with Equation (4.62). Here, $R_x(\Delta\varphi_x)$ denotes an operator that rotates the system by an angle $\Delta\varphi_x$ about the x-axis, etc. Suppose that the system is in an eigenstate of zero overall orbital angular momentum (i.e., an $l = 0$ state). We know that the system is also in an eigenstate of zero orbital angular momentum about any particular axis. This follows because $l = 0$ implies $m = 0$, according to the previous section, and we can choose the z-axis to point in any direction. Thus,

$$R_x(\Delta\varphi_x)\,|0,0\rangle = \exp(0)\,|0,0\rangle = |0,0\rangle, \tag{4.71}$$

$$R_y(\Delta\varphi_y)\,|0,0\rangle = \exp(0)\,|0,0\rangle = |0,0\rangle, \tag{4.72}$$

$$R_z(\Delta\varphi_z)\,|0,0\rangle = \exp(0)\,|0,0\rangle = |0,0\rangle. \tag{4.73}$$

Clearly, a zero angular momentum state is invariant to rotations about *any* axis. Such a state must possess a spherically symmetric wavefunction.

Note that a rotation about the x-axis does not commute with a rotation about the y-axis. In other words, if the system is rotated an angle $\Delta\varphi_x$ about the x-axis, and then $\Delta\varphi_y$ about the y-axis, it ends up in a different state to that obtained by rotating an angle $\Delta\varphi_y$ about the y-axis, and then $\Delta\varphi_x$ about the x-axis. In quantum mechanics, this implies that $R_y(\Delta\varphi_y)\,R_x(\Delta\varphi_x) \neq R_x(\Delta\varphi_x)\,R_y(\Delta\varphi_y)$, or $L_y\,L_x \neq L_x\,L_y$, [see Equations (4.68)–(4.70)]. Thus, the noncommuting nature of the angular momentum operators is a direct consequence of the fact that rotations do not commute.

4.4 Eigenfunctions of Orbital Angular Momentum

In Cartesian coordinates, the three components of orbital angular momentum can be written

$$L_x = -\mathrm{i}\,\hbar\left(y\,\frac{\partial}{\partial z} - z\,\frac{\partial}{\partial y}\right), \tag{4.74}$$

$$L_y = -\mathrm{i}\,\hbar\left(z\,\frac{\partial}{\partial x} - x\,\frac{\partial}{\partial z}\right), \tag{4.75}$$

$$L_z = -\mathrm{i}\,\hbar\left(x\,\frac{\partial}{\partial y} - y\,\frac{\partial}{\partial x}\right), \tag{4.76}$$

using the Schrödinger representation. Transforming to standard spherical polar coordinates,

$$x = r\,\sin\theta\,\cos\varphi, \tag{4.77}$$

$$y = r\,\sin\theta\,\sin\varphi, \tag{4.78}$$

$$z = r\,\cos\theta, \tag{4.79}$$

we obtain

$$L_x = \mathrm{i}\,\hbar\left(\sin\varphi\,\frac{\partial}{\partial\theta} + \cot\theta\cos\varphi\,\frac{\partial}{\partial\varphi}\right), \tag{4.80}$$

$$L_y = -\mathrm{i}\,\hbar\left(\cos\varphi\,\frac{\partial}{\partial\theta} - \cot\theta\sin\varphi\,\frac{\partial}{\partial\varphi}\right), \tag{4.81}$$

$$L_z = -\mathrm{i}\,\hbar\,\frac{\partial}{\partial\varphi}. \tag{4.82}$$

Note that Equation (4.82) accords with Equation (4.57). The shift operators $L^{\pm} = L_x \pm \mathrm{i}\,L_y$ become

$$L^{\pm} = \pm\hbar\,\exp(\pm\mathrm{i}\,\varphi)\left(\frac{\partial}{\partial\theta} \pm \mathrm{i}\,\cot\theta\,\frac{\partial}{\partial\varphi}\right). \tag{4.83}$$

Now,

$$L^2 = L_x^2 + L_y^2 + L_z^2 = L_z^2 + (L^+\,L^- + L^-\,L^+)/2, \tag{4.84}$$

so

$$L^2 = -\hbar^2\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\,\sin\theta\,\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right). \tag{4.85}$$

The eigenvalue problem for L^2 takes the form

$$L^2\,\psi = \lambda\,\hbar^2\,\psi, \tag{4.86}$$

where $\psi(r,\theta,\varphi)$ is the wavefunction, and λ is a number. Let us write

$$\psi(r,\theta,\varphi) = R(r)\,Y(\theta,\varphi). \tag{4.87}$$

Equation (4.86) reduces to

$$\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\,\sin\theta\,\frac{\partial}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right)Y+\lambda\,Y=0, \tag{4.88}$$

where use has been made of Equation (4.85). As is well-known, square integrable solutions to this equation only exist when λ takes the values $l\,(l+1)$, where l is an integer. These solutions are known as *spherical harmonics*, and can be written

$$Y_{l\,m}(\theta,\varphi)=\sqrt{\frac{2\,l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}\,(-1)^m\,\mathrm{e}^{\,\mathrm{i}\,m\,\varphi}\,P_{l\,m}(\cos\varphi), \tag{4.89}$$

where m is a positive integer lying in the range $0\leq m\leq l$. Here, $P_{l\,m}(\xi)$ is an *associated Legendre function* satisfying the equation

$$\frac{d}{d\xi}\left[(1-\xi^2)\,\frac{dP_{l\,m}}{d\xi}\right]-\frac{m^2}{1-\xi^2}\,P_{l\,m}+l\,(l+1)\,P_{l\,m}=0. \tag{4.90}$$

We define

$$Y_{l\,-m}=(-1)^m\,Y_{l\,m}^{*}, \tag{4.91}$$

which allows m to take the negative values $-l\leq m<0$. The spherical harmonics are *orthogonal* functions, and are properly normalized with respect to integration over the entire solid angle:

$$\int_0^{\pi}\int_0^{2\pi}d\theta\,d\varphi\,\sin\theta\,Y_{l\,m}^{*}(\theta,\varphi)\,Y_{l'\,m'}(\theta,\varphi)=\delta_{l\,l'}\,\delta_{m\,m'}. \tag{4.92}$$

The spherical harmonics also form a complete set for representing general functions of θ and φ.

By definition,

$$L^2\,Y_{l\,m}=l\,(l+1)\,\hbar^2\,Y_{l\,m}, \tag{4.93}$$

where l is an integer. It follows from Equations (4.82) and (4.89) that

$$L_z\,Y_{l\,m}=m\,\hbar\,Y_{l\,m}, \tag{4.94}$$

where m is an integer lying in the range $-l\leq m\leq l$. Thus, the wavefunction $\psi(r,\theta,\varphi)=R(r)\,Y_{l,m}(\theta,\phi)$, where R is a general function, has all of the expected features of the wavefunction of a simultaneous eigenstate of L^2 and L_z belonging to the quantum numbers l and m. The well-known formula

$$\frac{dP_{l\,m}}{d\xi}=-\frac{1}{\sqrt{1-\xi^2}}\,P_{l\,m+1}-\frac{m\,\xi}{1-\xi^2}\,P_{l\,m}=\frac{(l+m)\,(l-m+1)}{\sqrt{1-\xi^2}}\,P_{l\,m-1}+\frac{m\,\xi}{1-\xi^2}\,P_{l\,m} \tag{4.95}$$

can be combined with Equations (4.83) and (4.89) to give

$$L^{+}\,Y_{l\,m}=[l\,(l+1)-m\,(m+1)]^{1/2}\,\hbar\,Y_{l\,m+1}, \tag{4.96}$$

$$L^{-}\,Y_{l\,m}=[l\,(l+1)-m\,(m-1)]^{1/2}\,\hbar\,Y_{l\,m-1}. \tag{4.97}$$

These equations are equivalent to Equations (4.55)–(4.56). Note that a spherical harmonic wavefunction is symmetric about the z-axis (i.e., independent of φ) whenever $m = 0$, and is spherically symmetric whenever $l = 0$ (since $Y_{0\,0} = 1/\sqrt{4\pi}$).

In summary, by solving directly for the eigenfunctions of L^2 and L_z in the Schrödinger representation, we have been able to reproduce all of the results of Section 4.2. Nevertheless, the results of Section 4.2 are more general than those obtained in this section, because they still apply when the quantum number l takes on half-integer values.

4.5 Motion in Central Field

Consider a particle of mass M moving in a spherically symmetric potential. The Hamiltonian takes the form

$$H = \frac{p^2}{2\,M} + V(r). \tag{4.98}$$

Adopting Schrödinger's representation, we can write $\mathbf{p} = -(\mathrm{i}/\hbar)\nabla$. Hence,

$$H = -\frac{\hbar^2}{2\,M}\,\nabla^2 + V(r). \tag{4.99}$$

When written in spherical polar coordinates, the above equation becomes

$$H = -\frac{\hbar^2}{2\,M}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\,r^2\,\frac{\partial}{\partial r} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\,\sin\theta\,\frac{\partial}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] + V(r). \tag{4.100}$$

Comparing this equation with Equation (4.85), we find that

$$H = \frac{\hbar^2}{2\,M}\left(-\frac{1}{r^2}\frac{\partial}{\partial r}\,r^2\,\frac{\partial}{\partial r} + \frac{L^2}{\hbar^2 r^2}\right) + V(r). \tag{4.101}$$

Now, we know that the three components of angular momentum commute with L^2 (see Section 4.1). We also know, from Equations (4.80)–(4.82), that L_x, L_y, and L_z take the form of partial derivative operators involving only *angular* coordinates, when written in terms of spherical polar coordinates using the Schrödinger representation. It follows from Equation (4.101) that all three components of the angular momentum commute with the Hamiltonian:

$$[\mathbf{L}, H] = 0. \tag{4.102}$$

It is also easily seen that L^2 (which can be expressed as a purely angular differential operator) commutes with the Hamiltonian:

$$[L^2, H] = 0. \tag{4.103}$$

According to Section 3.2, the previous two equations ensure that the angular momentum $\mathbf{L}$ and its magnitude squared L^2 are both constants of the motion. This is as expected for a spherically symmetric potential.

Consider the energy eigenvalue problem

$$H\,\psi = E\,\psi, \tag{4.104}$$

where E is a number. Since L^2 and L_z commute with each other and the Hamiltonian, it is always possible to represent the state of the system in terms of the simultaneous eigenstates of L^2, L_z, and H. But, we already know that the most general form for the wavefunction of a simultaneous eigenstate of L^2 and L_z is (see previous section)

$$\psi(r,\theta,\varphi) = R(r)\,Y_{l\,m}(\theta,\varphi). \tag{4.105}$$

Substituting Equation (4.105) into Equation (4.101), and making use of Equation (4.93), we obtain

$$\left[\frac{\hbar^2}{2\,M}\left(-\frac{1}{r^2}\frac{d}{dr}\,r^2\,\frac{d}{dr}+\frac{l\,(l+1)}{r^2}\right)+V(r)-E\right]R = 0. \tag{4.106}$$

This is a *Sturm-Liouville equation* for the function $R(r)$. We know, from the general properties of this type of equation, that if $R(r)$ is required to be well-behaved at $r = 0$ and as $r \rightarrow \infty$ then solutions only exist for a discrete set of values of E. These are the energy eigenvalues. In general, the energy eigenvalues depend on the quantum number l, but are independent of the quantum number m.

4.6 Energy Levels of Hydrogen Atom

Consider a hydrogen atom, for which the potential takes the specific form

$$V(r) = -\frac{e^2}{4\pi\,\epsilon_0\,r}. \tag{4.107}$$

The radial eigenfunction $R(r)$ satisfies Equation (4.106), which can be written

$$\left[\frac{\hbar^2}{2\,\mu}\left(-\frac{1}{r^2}\frac{d}{dr}\,r^2\,\frac{d}{dr}+\frac{l\,(l+1)}{r^2}\right)-\frac{e^2}{4\pi\,\epsilon_0\,r}-E\right]R = 0. \tag{4.108}$$

Here, $\mu = m_e\,m_p/(m_e+m_p) \simeq m_e$ is the *reduced mass*, which takes into account the fact that the electron (of mass m_e) and the proton (of mass m_p) both rotate about a common centre, which is equivalent to a particle of mass μ rotating about a fixed point. Let us write the product $r\,R(r)$ as the function $P(r)$. The above equation transforms to

$$\frac{d^2P}{dr^2}-\frac{2\,\mu}{\hbar^2}\left[\frac{l\,(l+1)\,\hbar^2}{2\,\mu\,r^2}-\frac{e^2}{4\pi\,\epsilon_0\,r}-E\right]P = 0, \tag{4.109}$$

which is the one-dimensional Schrödinger equation for a particle of mass μ moving in the *effective potential*

$$V_{\rm eff}(r) = -\frac{e^2}{4\pi\,\epsilon_0\,r}+\frac{l\,(l+1)\,\hbar^2}{2\,\mu\,r^2}. \tag{4.110}$$

The effective potential has a simple physical interpretation. The first part is the attractive Coulomb potential, and the second part corresponds to the repulsive centrifugal force.

Let

$$a = \sqrt{\frac{-\hbar^2}{2\,\mu\,E}}, \tag{4.111}$$

and $y = r/a$, with

$$P(r) = f(y)\exp(-y). \tag{4.112}$$

Here, it is assumed that the energy eigenvalue E is negative. Equation (4.109) transforms to

$$\left(\frac{d^2}{dy^2} - 2\,\frac{d}{dy} - \frac{l\,(l+1)}{y^2} + \frac{2\,\mu\,e^2\,a}{4\pi\,\epsilon_0\,\hbar^2\,y}\right)f = 0. \tag{4.113}$$

Let us look for a power-law solution of the form

$$f(y) = \sum_n c_n\,y^n. \tag{4.114}$$

Substituting this solution into Equation (4.113), we obtain

$$\sum_n c_n\left[n\,(n-1)\,y^{n-2} - 2\,n\,y^{n-1} - l\,(l+1)\,y^{n-2} + \frac{2\,\mu\,e^2\,a}{4\pi\,\epsilon_0\,\hbar^2}\,y^{n-1}\right] = 0. \tag{4.115}$$

Equating the coefficients of y^{n-2} gives

$$c_n\,[n\,(n-1) - l\,(l+1)] = c_{n-1}\left[2\,(n-1) - \frac{2\,\mu\,e^2\,a}{4\pi\,\epsilon_0\,\hbar^2}\right]. \tag{4.116}$$

Now, the power law series (4.114) must terminate at small n, at some positive value of n, otherwise $f(y)$ would behave unphysically as $y \to 0$. This is only possible if $[n_{\rm min}(n_{\rm min} - 1) - l\,(l+1)] = 0$, where the first term in the series is $c_{n_{\rm min}}\,y^{n_{\rm min}}$. There are two possibilities: $n_{\rm min} = -l$ or $n_{\rm min} = l + 1$. The former predicts unphysical behavior of the wavefunction at $y = 0$. Thus, we conclude that $n_{\rm min} = l + 1$. Note that for an $l = 0$ state there is a finite probability of finding the electron at the nucleus, whereas for an $l > 0$ state there is zero probability of finding the electron at the nucleus (i.e., $|\psi|^2 = 0$ at $r = 0$, except when $l = 0$). Note, also, that it is only possible to obtain sensible behavior of the wavefunction as $r \to 0$ if l is an integer.

For large values of y, the ratio of successive terms in the series (4.114) is

$$\frac{c_n\,y}{c_{n-1}} = \frac{2\,y}{n}, \tag{4.117}$$

according to Equation (4.116). This is the same as the ratio of successive terms in the series

$$\sum_n \frac{(2\,y)^n}{n!}, \tag{4.118}$$

which converges to $\exp(2\,y)$. We conclude that $f(y) \rightarrow \exp(2\,y)$ as $y \rightarrow \infty$. It follows from Equation (4.112) that $R(r) \rightarrow \exp(r/a)/r$ as $r \rightarrow \infty$. This does not correspond to physically acceptable behavior of the wavefunction, since $\int d^3x'\,|\psi|^2$ must be finite. The only way in which we can avoid this unphysical behavior is if the series (4.114) terminates at some maximum value of n. According to the recursion relation (4.116), this is only possible if

$$\frac{\mu\,e^2\,a}{4\pi\,\epsilon_0\,\hbar^2} = n, \tag{4.119}$$

where the last term in the series is $c_n\,y^n$. It follows from Equation (4.111) that the energy eigenvalues are *quantized*, and can only take the values

$$E = \frac{E_0}{n^2}, \tag{4.120}$$

where

$$E_0 = -\frac{\mu\,e^4}{32\pi^2\,\epsilon_0^2\,\hbar^2} = -13.6\,\text{eV} \tag{4.121}$$

is the ground state energy. Here, n is a positive integer which must exceed the quantum number l, otherwise there would be no terms in the series (4.114).

The properly normalized wavefunction of a hydrogen atom is written

$$\psi(r,\theta,\varphi) = R_{n\,l}(r)\,Y_{l\,m}(\theta,\varphi), \tag{4.122}$$

where

$$R_{n\,l}(r) = \mathcal{R}_{n\,l}(r/a), \tag{4.123}$$

and

$$a = n\,a_0. \tag{4.124}$$

Here,

$$a_0 = \frac{4\pi\,\epsilon_0\,\hbar^2}{\mu\,e^2} = 5.3\times 10^{-11}\ \text{meters} \tag{4.125}$$

is the *Bohr radius*, and $\mathcal{R}_{n\,l}(x)$ is a well-behaved solution of the differential equation

$$\left[\frac{1}{x^2}\frac{d}{dx}\,x^2\,\frac{d}{dx} - \frac{l\,(l+1)}{x^2} + \frac{2\,n}{x} - 1\right]\mathcal{R}_{n\,l} = 0 \tag{4.126}$$

that is consistent with the normalization constraint

$$\int_0^\infty dr\,r^2\,[R_{n\,l}(r)]^2 = 1. \tag{4.127}$$

Finally, the $Y_{l\,m}$ are spherical harmonics. The restrictions on the quantum numbers are $|m| \leq l < n$, where n is a positive integer, l a non-negative integer, and m an integer.

The ground state of hydrogen corresponds to $n = 1$. The only permissible values of the other quantum numbers are $l = 0$ and $m = 0$. Thus, the ground state is a spherically symmetric, zero angular momentum state. The next energy level corresponds to $n = 2$. The other quantum numbers are allowed to take the values $l = 0$, $m = 0$ or $l = 1$, $m = -1, 0, 1$. Thus, there are $n = 2$ states with non-zero angular momentum. Note that the energy levels given in Equation (4.120) are independent of the quantum number l, despite the fact that l appears in the radial eigenfunction equation (4.126). This is a special property of a $1/r$ Coulomb potential.

In addition to the quantized negative energy states of the hydrogen atom, which we have just found, there is also a continuum of unbound positive energy states.

Exercises

4.1 Demonstrate directly from the fundamental commutation relations for angular momentum, (4.11), that $[L^2, L_z] = 0$, $[L^\pm, L_z] = \mp\hbar\, L^\pm$, and $[L^+, L^-] = 2\,\hbar\, L_z$.

4.2 Demonstrate from Equations (4.74)–(4.79) that

$$L_x = \mathrm{i}\,\hbar\left(\sin\varphi\,\frac{\partial}{\partial\theta} + \cot\theta\cos\varphi\,\frac{\partial}{\partial\varphi}\right),$$

$$L_y = -\mathrm{i}\,\hbar\left(\cos\varphi\,\frac{\partial}{\partial\theta} - \cot\theta\sin\varphi\,\frac{\partial}{\partial\varphi}\right),$$

$$L_z = -\mathrm{i}\,\hbar\,\frac{\partial}{\partial\varphi},$$

where θ, φ are conventional spherical polar angles.

4.3 A system is in the state $\psi(\theta,\varphi) = Y_{l\,m}(\theta,\varphi)$. Evaluate $\langle L_x\rangle$, $\langle L_y\rangle$, $\langle L_x^2\rangle$, and $\langle L_y^2\rangle$.

4.4 Derive Equations (4.96) and (4.97) from Equation (4.95).

4.5 Find the eigenvalues and eigenfunctions (in terms of the angles θ and φ) of L_x.

4.6 Consider a beam of particles with $l = 1$. A measurement of L_x yields the result $\hbar$. What values will be obtained by a subsequent measurement of L_z, and with what probabilities? Repeat the calculation for the cases in which the measurement of L_x yields the results 0 and $-\hbar$.

4.7 The Hamiltonian for an axially symmetric rotator is given by

$$H = \frac{L_x^2 + L_y^2}{2\,I_1} + \frac{L_z^2}{2\,I_2}.$$

What are the eigenvalues of H?

4.8 The expectation value of $f(\mathbf{x},\mathbf{p})$ in any stationary state is a constant. Calculate

$$0 = \frac{d}{dt}\left(\langle\mathbf{x}\cdot\mathbf{p}\rangle\right) = \frac{\mathrm{i}}{\hbar}\,\langle[H, \mathbf{x}\cdot\mathbf{p}]\rangle$$

for a Hamiltonian of the form

$$H = \frac{p^2}{2\,m} + V(r).$$

Hence, show that

$$\left\langle\frac{p^2}{2\,m}\right\rangle = \frac{1}{2}\left\langle r\,\frac{dV}{dr}\right\rangle$$

in a stationary state. This is another form of the *Virial theorem*. (See Exercise 3.8.)

4.9 Use the Virial theorem of the previous exercise to prove that

$$\left\langle\frac{1}{r}\right\rangle = \frac{1}{n^2\,a_0}$$

for an energy eigenstate of the hydrogen atom.

4.10 Demonstrate that the first few properly normalized radial wavefunctions of the hydrogen atom take the form:

(a)

$$R_{1\,0}(r) = \frac{2}{a_0^{3/2}}\,\exp\left(-\frac{r}{a_0}\right).$$

(b)

$$R_{2\,0}(r) = \frac{2}{(2\,a_0)^{3/2}}\left(1 - \frac{r}{2\,a_0}\right)\exp\left(-\frac{r}{2\,a_0}\right).$$

(c)

$$R_{2\,1}(r) = \frac{1}{\sqrt{3}\,(2\,a_0)^{3/2}}\,\frac{r}{a_0}\,\exp\left(-\frac{r}{2\,a_0}\right).$$

5 Spin Angular Momentum

5.1 Introduction

Up to now, we have tacitly assumed that the state of a particle in quantum mechanics can be completely specified by giving the wavefunction ψ as a function of the spatial coordinates x, y, and z. Unfortunately, there is a wealth of experimental evidence that suggests that this simplistic approach is incomplete.

Consider an isolated system at rest, and let the eigenvalue of its total angular momentum be $j\,(j+1)\,\hbar^2$. According to the theory of orbital angular momentum outlined in Sections 4.4 and 4.5, there are two possibilities. For a system consisting of a single particle, $j=0$. For a system consisting of two (or more) particles, j is a non-negative integer. However, this does not agree with observations, because we often encounter systems that appear to be structureless, and yet have $j\neq 0$. Even worse, systems where j has half-integer values abound in nature. In order to explain this apparent discrepancy between theory and experiments, Gouldsmit and Uhlenbeck (in 1925) introduced the concept of an internal, purely quantum mechanical, angular momentum called *spin*. For a particle with spin, the total angular momentum in the rest frame is non-vanishing.

5.2 Properties of Spin Angular Momentum

Let us denote the three components of the spin angular momentum of a particle by the Hermitian operators $(S_x, S_y, S_z)\equiv \mathbf{S}$. We assume that these operators obey the fundamental commutation relations (4.8)–(4.10) for the components of an angular momentum. Thus, we can write

$$\mathbf{S}\times\mathbf{S} = {\rm i}\,\hbar\,\mathbf{S}. \tag{5.1}$$

We can also define the operator

$$S^2 = S_x^{\,2}+S_y^{\,2}+S_z^{\,2}. \tag{5.2}$$

According to the quite general analysis of Section 4.1,

$$[\mathbf{S}, S^2] = 0. \tag{5.3}$$

Thus, it is possible to find simultaneous eigenstates of S^2 and S_z. These are denoted $|s, s_z\rangle$, where

$$S_z\,|s, s_z\rangle = s_z\,\hbar\,|s, s_z\rangle, \tag{5.4}$$

$$S^2\,|s, s_z\rangle = s\,(s+1)\,\hbar^2\,|s, s_z\rangle. \tag{5.5}$$

According to the equally general analysis of Section 4.2, the quantum number s can, in principle, take integer or half-integer values, and the quantum number s_z can only take the values $s, s-1\cdots -s+1, -s$.

Spin angular momentum clearly has many properties in common with orbital angular momentum. However, there is one vitally important difference. Spin angular momentum operators *cannot* be expressed in terms of position and momentum operators, like in Equations (4.1)–(4.3), because this identification depends on an analogy with classical mechanics, and the concept of spin is purely quantum mechanical: i.e., it has no analogy in classical physics. Consequently, the restriction that the quantum number of the overall angular momentum must take *integer* values is lifted for spin angular momentum, since this restriction (found in Sections 4.3 and 4.4) depends on Equations (4.1)–(4.3). In other words, the spin quantum number s is allowed to take *half-integer* values.

Consider a spin one-half particle, for which

$$S_z\,|\pm\rangle = \pm\frac{\hbar}{2}\,|\pm\rangle, \tag{5.6}$$

$$S^2\,|\pm\rangle = \frac{3\,\hbar^2}{4}\,|\pm\rangle. \tag{5.7}$$

Here, the $|\pm\rangle$ denote eigenkets of the S_z operator corresponding to the eigenvalues $\pm\hbar/2$. These kets are mutually orthogonal (since S_z is an Hermitian operator), so

$$\langle +|-\rangle = 0. \tag{5.8}$$

They are also properly normalized and complete, so that

$$\langle +|+\rangle = \langle -|-\rangle = 1, \tag{5.9}$$

and

$$|+\rangle\langle +| + |-\rangle\langle -| = 1. \tag{5.10}$$

It is easily verified that the Hermitian operators defined by

$$S_x = \frac{\hbar}{2}\,(\,|+\rangle\langle -| + |-\rangle\langle +|\,), \tag{5.11}$$

$$S_y = \frac{\mathrm{i}\,\hbar}{2}\,(\,-|+\rangle\langle -| + |-\rangle\langle +|\,), \tag{5.12}$$

$$S_z = \frac{\hbar}{2}\,(\,|+\rangle\langle +| - |-\rangle\langle -|\,), \tag{5.13}$$

satisfy the commutation relations (4.8)–(4.10) (with the L_j replaced by the S_j). The operator S^2 takes the form

$$S^2 = \frac{3\,\hbar^2}{4}. \tag{5.14}$$

It is also easily demonstrated that S^2 and S_z, defined in this manner, satisfy the eigenvalue relations (5.6)–(5.7). Equations (5.11)–(5.14) constitute a realization of the spin operators $\mathbf{S}$ and S^2 (for a spin one-half particle) in *spin space* (i.e., the Hilbert sub-space consisting of kets which correspond to the different spin states of the particle).

5.3 Wavefunction of Spin One-Half Particle

The state of a spin one-half particle is represented as a vector in ket space. Let us suppose that this space is spanned by the basis kets $|x', y', z', \pm\rangle$. Here, $|x', y', z', \pm\rangle$ denotes a simultaneous eigenstate of the position operators x, y, z, and the spin operator S_z, corresponding to the eigenvalues x', y', z', and $\pm\hbar/2$, respectively. The basis kets are assumed to satisfy the completeness relation

$$\int\!\!\int\!\!\int dx'dy'dz' \left(|x', y', z', +\rangle\langle x', y', z', +| + |x', y', z', -\rangle\langle x', y', z', -|\right) = 1. \tag{5.15}$$

It is helpful to think of the ket $|x', y', z', +\rangle$ as the product of two kets—a position space ket $|x', y', z'\rangle$, and a spin space ket $|+\rangle$. We assume that such a product obeys the commutative and distributive axioms of multiplication:

$$|x', y', z'\rangle|+\rangle = |+\rangle|x', y', z'\rangle, \tag{5.16}$$

$$(c'\,|x', y', z'\rangle + c''\,|x'', y'', z''\rangle)\,|+\rangle = c'\,|x', y', z'\rangle|+\rangle + c''\,|x'', y'', z''\rangle|+\rangle, \tag{5.17}$$

$$|x', y', z'\rangle\,(c_+\,|+\rangle + c_-\,|-\rangle) = c_+\,|x', y', z'\rangle|+\rangle + c_-\,|x', y', z'\rangle|-\rangle, \tag{5.18}$$

where the c's are numbers. We can give meaning to any position space operator (such as L_z) acting on the product $|x', y', z'\rangle|+\rangle$ by assuming that it operates only on the $|x', y', z'\rangle$ factor, and commutes with the $|+\rangle$ factor. Similarly, we can give a meaning to any spin operator (such as S_z) acting on $|x', y', z'\rangle|+\rangle$ by assuming that it operates only on $|+\rangle$, and commutes with $|x', y', z'\rangle$. This implies that every position space operator commutes with every spin operator. In this manner, we can give meaning to the equation

$$|x', y', z', \pm\rangle = |x', y', z'\rangle|\pm\rangle = |\pm\rangle|x', y', z'\rangle. \tag{5.19}$$

The multiplication in the above equation is of a quite different type to any that we have encountered previously. The ket vectors $|x', y', z'\rangle$ and $|\pm\rangle$ lie in two completely separate vector spaces, and their product $|x', y', z'\rangle|\pm\rangle$ lies in a third vector space. In mathematics, the latter space is termed the *product space* of the former spaces, which are termed *factor spaces*. The number of dimensions of a product space is equal to the product of the number of dimensions of each of the factor spaces. A general ket of the product space is not of the form (5.19), but is instead a sum or integral of kets of this form.

A general state A of a spin one-half particle is represented as a ket $||A\rangle\rangle$ in the product of the spin and position spaces. This state can be completely specified by *two* wavefunctions:

$$\psi_+(x', y', z') = \langle x', y', z'|\langle +||A\rangle\rangle, \tag{5.20}$$

$$\psi_-(x', y', z') = \langle x', y', z'|\langle -||A\rangle\rangle. \tag{5.21}$$

The probability of observing the particle in the region x' to $x' + dx'$, y' to $y' + dy'$, and z' to $z' + dz'$, with $s_z = +1/2$ is $|\psi_+(x', y', z')|^2\,dx'dy'dz'$. Likewise, the probability of observing the particle in the region x' to $x' + dx'$, y' to $y' + dy'$, and z' to $z' + dz'$, with $s_z = -1/2$ is $|\psi_-(x', y', z')|^2\,dx'dy'dz'$. The normalization condition for the wavefunctions is

$$\int\!\!\int\!\!\int dx'dy'dz' \left(|\psi_+|^2 + |\psi_-|^2\right) = 1. \tag{5.22}$$

5.4 Rotation Operators in Spin Space

Let us, for the moment, forget about the spatial position of the particle, and concentrate on its spin state. A general spin state A is represented by the ket

$$|A\rangle = \langle +|A\rangle|+\rangle + \langle -|A\rangle|-\rangle \tag{5.23}$$

in spin space. In Section 4.3, we were able to construct an operator $R_z(\Delta\varphi)$ that rotates the system through an angle $\Delta\varphi$ about the z-axis in position space. Can we also construct an operator $T_z(\Delta\varphi)$ that rotates the system through an angle $\Delta\varphi$ about the z-axis in spin space? By analogy with Equation (4.62), we would expect such an operator to take the form

$$T_z(\Delta\varphi) = \exp(-\mathrm{i}\, S_z\, \Delta\varphi/\hbar). \tag{5.24}$$

Thus, after rotation, the ket $|A\rangle$ becomes

$$|A_R\rangle = T_z(\Delta\varphi)\,|A\rangle. \tag{5.25}$$

To demonstrate that the operator (5.24) really does rotate the spin of the system, let us consider its effect on $\langle S_x\rangle$. Under rotation, this expectation value changes as follows:

$$\langle S_x\rangle \rightarrow \langle A_R|\,S_x\,|A_R\rangle = \langle A|\,T_z^{\dagger}\,S_x\,T_z\,|A\rangle. \tag{5.26}$$

Thus, we need to compute

$$\exp(\,\mathrm{i}\,S_z\,\Delta\varphi/\hbar)\,S_x\,\exp(-\mathrm{i}\,S_z\,\Delta\varphi/\hbar). \tag{5.27}$$

This can be achieved in two different ways.

First, we can use the explicit formula for S_x given in Equation (5.11). We find that Equation (5.27) becomes

$$\frac{\hbar}{2}\,\exp(\,\mathrm{i}\,S_z\,\Delta\varphi/\hbar)\,(\,|+\rangle\langle -| + |-\rangle\langle +|\,)\,\exp(-\mathrm{i}\,S_z\,\Delta\varphi/\hbar), \tag{5.28}$$

or

$$\frac{\hbar}{2}\left(\mathrm{e}^{\,\mathrm{i}\,\Delta\varphi/2}\,|+\rangle\langle -|\,\mathrm{e}^{\,\mathrm{i}\,\Delta\varphi/2} + \mathrm{e}^{\,-\mathrm{i}\,\Delta\varphi/2}\,|-\rangle\langle +|\,\mathrm{e}^{\,-\mathrm{i}\,\Delta\varphi/2}\right), \tag{5.29}$$

which reduces to

$$S_x\,\cos\Delta\varphi - S_y\,\sin\Delta\varphi, \tag{5.30}$$

where use has been made of Equations (5.11)–(5.13).

A second approach is to use the so called *Baker-Hausdorff lemma*. This takes the form

$$\exp(\,\mathrm{i}\,G\,\lambda)\,A\,\exp(-\mathrm{i}\,G\,\lambda) = A + \mathrm{i}\,\lambda\,[G,A] + \left(\frac{\mathrm{i}^2\lambda^2}{2!}\right)[G,[G,A]] + \\ + \left(\frac{\mathrm{i}^3\lambda^3}{3!}\right)[G,[G,[G,A]]] + \cdots, \tag{5.31}$$

where G is an Hermitian operator, and λ a real parameter. The proof of this lemma is left as an exercise. Applying the Baker-Hausdorff lemma to Equation (5.27), we obtain

$$S_x + \left(\frac{\mathrm{i}\,\varDelta\varphi}{\hbar}\right)[S_z, S_x] + \left(\frac{1}{2!}\right)\left(\frac{\mathrm{i}\,\varDelta\varphi}{\hbar}\right)^2 [S_z, [S_z, S_x]] + \cdots, \tag{5.32}$$

which reduces to

$$S_x\left[1 - \frac{(\varDelta\varphi)^2}{2!} + \frac{(\varDelta\varphi)^4}{4!} + \cdots\right] - S_y\left[\varphi - \frac{(\varDelta\varphi)^3}{3!} + \frac{(\varDelta\varphi)^5}{5!} + \cdots\right], \tag{5.33}$$

or

$$S_x \cos\varDelta\varphi - S_y \sin\varDelta\varphi, \tag{5.34}$$

where use has been made of Equation (5.1). The second proof is more general than the first, because it only uses the fundamental commutation relation (5.1), and is, therefore, valid for systems with spin angular momentum higher than one-half.

For a spin one-half system, both methods imply that

$$\langle S_x\rangle \rightarrow \langle S_x\rangle \cos\varDelta\varphi - \langle S_y\rangle \sin\varDelta\varphi \tag{5.35}$$

under the action of the rotation operator (5.24). It is straightforward to show that

$$\langle S_y\rangle \rightarrow \langle S_y\rangle \cos\varDelta\varphi + \langle S_x\rangle \sin\varDelta\varphi. \tag{5.36}$$

Furthermore,

$$\langle S_z\rangle \rightarrow \langle S_z\rangle, \tag{5.37}$$

because S_z commutes with the rotation operator. Equations (5.35)–(5.37) demonstrate that the operator (5.24) rotates the expectation value of $\mathbf{S}$ by an angle $\varDelta\varphi$ about the z-axis. In fact, the expectation value of the spin operator behaves like a classical vector under rotation:

$$\langle S_k\rangle \rightarrow \sum_l R_{k\,l}\,\langle S_l\rangle, \tag{5.38}$$

where the $R_{k\,l}$ are the elements of the conventional rotation matrix for the rotation in question. It is clear, from our second derivation of the result (5.35), that this property is not restricted to the spin operators of a spin one-half system. In fact, we have effectively demonstrated that

$$\langle J_k\rangle \rightarrow \sum_l R_{k\,l}\,\langle J_l\rangle, \tag{5.39}$$

where the J_k are the generators of rotation, satisfying the fundamental commutation relation $\mathbf{J}\times\mathbf{J} = \mathrm{i}\,\hbar\,\mathbf{J}$, and the rotation operator about the kth axis is written $R_k(\varDelta\varphi) = \exp(-\mathrm{i}\,J_k\,\varDelta\varphi/\hbar)$.

Consider the effect of the rotation operator (5.24) on the state ket (5.23). It is easily seen that

$$T_z(\varDelta\varphi)\,|A\rangle = \mathrm{e}^{-\mathrm{i}\,\varDelta\varphi/2}\,\langle +|A\rangle|+\rangle + \mathrm{e}^{\,\mathrm{i}\,\varDelta\varphi/2}\,\langle -|A\rangle|-\rangle. \tag{5.40}$$

Consider a rotation by 2π radians. We find that

$$|A\rangle \rightarrow T_z(2\pi)\,|A\rangle = -|A\rangle. \tag{5.41}$$

Note that a ket rotated by 2π radians differs from the original ket by a *minus* sign. In fact, a rotation by 4π radians is needed to transform a ket into itself. The minus sign does not affect the expectation value of $\mathbf{S}$, since $\mathbf{S}$ is sandwiched between $\langle A|$ and $|A\rangle$, both of which change sign. Nevertheless, the minus sign does give rise to observable consequences, as we shall see presently.

5.5 Magnetic Moments

Consider a particle of electric charge q and speed v performing a circular orbit of radius r in the x-y plane. The charge is equivalent to a current loop of radius r in the x-y plane carrying current $I = q\,v/2\pi\,r$. The magnetic moment $\boldsymbol{\mu}$ of the loop is of magnitude $\pi\,r^2\,I$ and is directed along the z-axis. Thus, we can write

$$\boldsymbol{\mu} = \frac{q}{2}\,\mathbf{x}\times\mathbf{v}, \tag{5.42}$$

where $\mathbf{x}$ and $\mathbf{v}$ are the vector position and velocity of the particle, respectively. However, we know that $\mathbf{p} = \mathbf{v}/m$, where $\mathbf{p}$ is the vector momentum of the particle, and m is its mass. We also know that $\mathbf{L} = \mathbf{x}\times\mathbf{p}$, where $\mathbf{L}$ is the orbital angular momentum. It follows that

$$\boldsymbol{\mu} = \frac{q}{2\,m}\,\mathbf{L}. \tag{5.43}$$

Using the usual analogy between classical and quantum mechanics, we expect the above relation to also hold between the quantum mechanical operators, $\boldsymbol{\mu}$ and $\mathbf{L}$, which represent magnetic moment and orbital angular momentum, respectively. This is indeed found to the the case.

Spin angular momentum also gives rise to a contribution to the magnetic moment of a charged particle. In fact, relativistic quantum mechanics predicts that a charged particle possessing spin must also possess a corresponding magnetic moment (this was first demonstrated by Dirac—see Chapter 11). We can write

$$\boldsymbol{\mu} = \frac{q}{2\,m}\,(\mathbf{L} + g\,\mathbf{S}), \tag{5.44}$$

where g is called the *gyromagnetic ratio*. For an electron this ratio is found to be

$$g_e = 2\left(1 + \frac{1}{2\pi}\frac{e^2}{4\pi\,\epsilon_0\,\hbar\,c}\right). \tag{5.45}$$

The factor 2 is correctly predicted by Dirac's relativistic theory of the electron (see Chapter 11). The small correction $1/(2\pi\,137)$, derived originally by Schwinger, is due to quantum field effects. We shall ignore this correction in the following, so

$$\boldsymbol{\mu} \simeq -\frac{e}{2\,m_e}\,(\mathbf{L} + 2\,\mathbf{S}) \tag{5.46}$$

for an electron (here, $e > 0$).

5.6 Spin Precession

The Hamiltonian for an electron at rest in a z-directed magnetic field, $\mathbf{B} = B\,\mathbf{e}_z$, is

$$H = -\boldsymbol{\mu}\cdot\mathbf{B} = \left(\frac{e}{m_e}\right)\mathbf{S}\cdot\mathbf{B} = \omega\,S_z, \tag{5.47}$$

where

$$\omega = \frac{e\,B}{m_e}. \tag{5.48}$$

According to Equation (3.28), the time evolution operator for this system is

$$T(t,0) = \exp(-\mathrm{i}\,H\,t/\hbar) = \exp(-\mathrm{i}\,S_z\,\omega\,t/\hbar). \tag{5.49}$$

It can be seen, by comparison with Equation (5.24), that the time evolution operator is precisely the same as the rotation operator for spin, with $\Delta\varphi$ set equal to $\omega\,t$. It is immediately clear that the Hamiltonian (5.47) causes the electron spin to precess about the z-axis with angular frequency ω. In fact, Equations (5.35)–(5.37) imply that

$$\langle S_x\rangle_t = \langle S_x\rangle_{t=0}\cos(\omega\,t) - \langle S_y\rangle_{t=0}\sin(\omega\,t), \tag{5.50}$$

$$\langle S_y\rangle_t = \langle S_y\rangle_{t=0}\cos(\omega\,t) + \langle S_x\rangle_{t=0}\sin(\omega\,t), \tag{5.51}$$

$$\langle S_z\rangle_t = \langle S_z\rangle_{t=0}. \tag{5.52}$$

The time evolution of the state ket is given by analogy with Equation (5.40):

$$|A,t\rangle = \mathrm{e}^{-\mathrm{i}\,\omega\,t/2}\,\langle +|A,0\rangle|+\rangle + \mathrm{e}^{\mathrm{i}\,\omega\,t/2}\,\langle -|A,0\rangle|-\rangle. \tag{5.53}$$

Note that it takes time $t = 4\pi/\omega$ for the state ket to return to its original state. By contrast, it only takes times $t = 2\pi/\omega$ for the spin vector to point in its original direction.

We now describe an experiment to detect the minus sign in Equation (5.41). An almost monoenergetic beam of neutrons is split in two, sent along two different paths, A and B, and then recombined. Path A goes through a magnetic field free region. However, path B enters a small region where a static magnetic field is present. As a result, a neutron state ket going along path B acquires a phase-shift $\exp(\mp\mathrm{i}\,\omega\,T/2)$ (the $\mp$ signs correspond to $s_z = \pm 1/2$ states). Here, T is the time spent in the magnetic field, and ω is the spin precession frequency

$$\omega = \frac{g_n\,e\,B}{m_p}. \tag{5.54}$$

This frequency is defined in an analogous manner to Equation (5.48). The gyromagnetic ratio for a neutron is found experimentally to be $g_n = -1.91$. (The magnetic moment of a neutron is entirely a quantum field effect). When neutrons from path A and path B meet they undergo interference. We expect the observed neutron intensity in the interference region to exhibit a $\cos(\pm\omega\,T/2 + \delta)$ variation, where δ is the phase difference between paths A and B in the absence of a magnetic field. In experiments, the time of flight T through the magnetic field region is kept constant, while the field-strength B is varied. It follows that the change in magnetic field required to produce successive maxima is

$$\Delta B = \frac{4\pi\,\hbar}{e\,g_n\,\lambdabar\,l}, \tag{5.55}$$

where l is the path-length through the magnetic field region, and $\lambdabar$ is the de Broglie wavelength over 2π of the neutrons. The above prediction has been verified experimentally to within a fraction of a percent. This prediction depends crucially on the fact that it takes a 4π rotation to return a state ket to its original state. If it only took a 2π rotation then ΔB would be half of the value given above, which does not agree with the experimental data.

5.7 Pauli Two-Component Formalism

We have seen, in Section 4.4, that the eigenstates of orbital angular momentum can be conveniently represented as spherical harmonics. In this representation, the orbital angular momentum operators take the form of differential operators involving only angular coordinates. It is conventional to represent the eigenstates of spin angular momentum as column (or row) matrices. In this representation, the spin angular momentum operators take the form of matrices.

The matrix representation of a spin one-half system was introduced by Pauli in 1926. Recall, from Section 5.4, that a general spin ket can be expressed as a linear combination of the two eigenkets of S_z belonging to the eigenvalues $\pm\hbar/2$. These are denoted $|\pm\rangle$. Let us represent these basis eigenkets as column vectors:

$$|+\rangle \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv \chi_+, \tag{5.56}$$

$$|-\rangle \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv \chi_-. \tag{5.57}$$

The corresponding eigenbras are represented as row vectors:

$$\langle +| \rightarrow (1,0) \equiv \chi_+^\dagger, \tag{5.58}$$

$$\langle -| \rightarrow (0,1) \equiv \chi_-^\dagger. \tag{5.59}$$

In this scheme, a general ket takes the form

$$|A\rangle = \langle +|A\rangle|+\rangle + \langle -|A\rangle|-\rangle \rightarrow \begin{pmatrix} \langle +|A\rangle \\ \langle -|A\rangle \end{pmatrix}, \tag{5.60}$$

and a general bra becomes

$$\langle A| = \langle A|+\rangle\langle +| + \langle A|-\rangle\langle -| \rightarrow (\langle A|+\rangle, \langle A|-\rangle). \tag{5.61}$$

The column vector (5.60) is called a two-component *spinor*, and can be written

$$\chi \equiv \begin{pmatrix} \langle +|A\rangle \\ \langle -|A\rangle \end{pmatrix} = \begin{pmatrix} c_+ \\ c_- \end{pmatrix} = c_+\chi_+ + c_-\chi_-, \tag{5.62}$$

where the $c_\pm$ are complex numbers. The row vector (5.61) becomes

$$\chi^\dagger \equiv (\langle A|+\rangle, \langle A|-\rangle) = (c_+^*, c_-^*) = c_+^*\,\chi_+^\dagger + c_-^*\,\chi_-^\dagger. \tag{5.63}$$

Consider the ket obtained by the action of a spin operator on ket A:

$$|A'\rangle = S_k\,|A\rangle. \tag{5.64}$$

This ket is represented as

$$|A'\rangle \rightarrow \begin{pmatrix} \langle +|A'\rangle \\ \langle -|A'\rangle \end{pmatrix} \equiv \chi'. \tag{5.65}$$

However,

$$\langle +|A'\rangle = \langle +|\,S_k\,|+\rangle\langle +|A\rangle + \langle +|\,S_k\,|-\rangle\langle -|A\rangle, \tag{5.66}$$

$$\langle -|A'\rangle = \langle -|\,S_k\,|+\rangle\langle +|A\rangle + \langle -|\,S_k\,|-\rangle\langle -|A\rangle, \tag{5.67}$$

or

$$\begin{pmatrix} \langle +|A'\rangle \\ \langle -|A'\rangle \end{pmatrix} = \begin{pmatrix} \langle +|\,S_k\,|+\rangle & \langle +|\,S_k\,|-\rangle \\ \langle -|\,S_k\,|+\rangle & \langle -|\,S_k\,|-\rangle \end{pmatrix} \begin{pmatrix} \langle +|A\rangle \\ \langle -|A\rangle \end{pmatrix}. \tag{5.68}$$

It follows that we can represent the operator/ket relation (5.64) as the matrix relation

$$\chi' = \left(\frac{\hbar}{2}\right)\sigma_k\,\chi, \tag{5.69}$$

where the σ_k are the matrices of the $\langle\pm|\,S_k\,|\pm\rangle$ values divided by $\hbar/2$. These matrices, which are called the *Pauli matrices*, can easily be evaluated using the explicit forms for the spin operators given in Equations (5.11)–(5.13). We find that

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{5.70}$$

$$\sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \tag{5.71}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.72}$$

Here, 1, 2, and 3 refer to x, y, and z, respectively. Note that, in this scheme, we are effectively representing the spin operators in terms of the Pauli matrices:

$$S_k \rightarrow \left(\frac{\hbar}{2}\right)\sigma_k. \tag{5.73}$$

The expectation value of S_k can be written in terms of spinors and the Pauli matrices:

$$\langle S_k\rangle = \langle A|\,S_k\,|A\rangle = \sum_{\pm}\langle A|\pm\rangle\langle\pm|\,S_k\,|\pm\rangle\langle\pm|A\rangle = \left(\frac{\hbar}{2}\right)\chi^\dagger\,\sigma_k\,\chi. \tag{5.74}$$

The fundamental commutation relation for angular momentum, Equation (5.1), can be combined with (5.73) to give the following commutation relation for the Pauli matrices:

$$\boldsymbol{\sigma}\times\boldsymbol{\sigma} = 2\,\mathrm{i}\,\boldsymbol{\sigma}. \tag{5.75}$$

It is easily seen that the matrices (5.70)–(5.72) actually satisfy these relations (i.e., $\sigma_1\,\sigma_2-\sigma_2\,\sigma_1 = 2\,\mathrm{i}\,\sigma_3$, plus all cyclic permutations). It is also easily seen that the Pauli matrices satisfy the anti-commutation relations

$$\{\sigma_i, \sigma_j\} = 2\,\delta_{ij}. \tag{5.76}$$

Here, $\{a, b\} \equiv a\,b + b\,a$.

Let us examine how the Pauli scheme can be extended to take into account the position of a spin one-half particle. Recall, from Section 5.3, that we can represent a general basis ket as the product of basis kets in position space and spin space:

$$|x', y', z', \pm\rangle = |x', y', z'\rangle|\pm\rangle = |\pm\rangle|x', y', z'\rangle. \tag{5.77}$$

The ket corresponding to state A is denoted $||A\rangle\rangle$, and resides in the product space of the position and spin ket spaces. State A is completely specified by the two wavefunctions

$$\psi_+(x', y', z') = \langle x', y', z'|\langle +||A\rangle\rangle, \tag{5.78}$$

$$\psi_-(x', y', z') = \langle x', y', z'|\langle -||A\rangle\rangle. \tag{5.79}$$

Consider the operator relation

$$||A'\rangle\rangle = S_k\,||A\rangle\rangle. \tag{5.80}$$

It is easily seen that

$$\langle x', y', z'|\langle +|A'\rangle\rangle = \langle +|\,S_k\,|+\rangle\langle x', y', z'|\langle +||A\rangle\rangle + \langle +|\,S_k\,|-\rangle\langle x', y', z'|\langle -||A\rangle\rangle, \tag{5.81}$$

$$\langle x', y', z'|\langle -|A'\rangle\rangle = \langle -|\,S_k\,|+\rangle\langle x', y', z'|\langle +||A\rangle\rangle + \langle -|\,S_k\,|-\rangle\langle x', y', z'|\langle -||A\rangle\rangle, \tag{5.82}$$

where use has been made of the fact that the spin operator S_k commutes with the eigenbras $\langle x', y', z'|$. It is fairly obvious that we can represent the operator relation (5.80) as a matrix relation if we generalize our definition of a spinor by writing

$$||A\rangle\rangle \rightarrow \begin{pmatrix} \psi_+(\mathbf{x}') \\ \psi_-(\mathbf{x}') \end{pmatrix} \equiv \chi, \tag{5.83}$$

and so on. The components of a spinor are now wavefunctions, instead of complex numbers. In this scheme, the operator equation (5.80) becomes simply

$$\chi' = \left(\frac{\hbar}{2}\right)\sigma_k\,\chi. \tag{5.84}$$

Consider the operator relation

$$||A'\rangle\rangle = p_k\,||A\rangle\rangle. \tag{5.85}$$

In the Schrödinger representation, we have

$$\langle x', y', z'|\langle +|A'\rangle\rangle = \langle x', y', z'|\, p_k \langle +||A\rangle\rangle = -\mathrm{i}\,\hbar \frac{\partial}{\partial x_k'} \langle x', y', z'|\langle +||A\rangle\rangle, \tag{5.86}$$

$$\langle x', y', z'|\langle -|A'\rangle\rangle = \langle x', y', z'|\, p_k \langle -||A\rangle\rangle = -\mathrm{i}\,\hbar \frac{\partial}{\partial x_k'} \langle x', y', z'|\langle -||A\rangle\rangle, \tag{5.87}$$

where use has been made of Equation (2.78). The above equation reduces to

$$\begin{pmatrix} \psi_+'(\mathbf{x}') \\ \psi_-'(\mathbf{x}') \end{pmatrix} = \begin{pmatrix} -\mathrm{i}\,\hbar\, \partial\psi_+(\mathbf{x}')/\partial x_k' \\ -\mathrm{i}\,\hbar\, \partial\psi_-(\mathbf{x}')/\partial x_k' \end{pmatrix}. \tag{5.88}$$

Thus, the operator equation (5.85) can be written

$$\chi' = p_k\, \chi, \tag{5.89}$$

where

$$p_k \rightarrow -\mathrm{i}\,\hbar \frac{\partial}{\partial x_k'}\, \mathbf{1}. \tag{5.90}$$

Here, $\mathbf{1}$ is the 2×2 unit matrix. In fact, any position operator (e.g., p_k or L_k) is represented in the Pauli scheme as some differential operator of the position eigenvalues multiplied by the 2×2 unit matrix.

What about combinations of position and spin operators? The most commonly occurring combination is a dot product: e.g., $\mathbf{S} \cdot \mathbf{L} = (\hbar/2)\, \boldsymbol{\sigma} \cdot \mathbf{L}$. Consider the hybrid operator $\boldsymbol{\sigma} \cdot \mathbf{a}$, where $\mathbf{a} \equiv (a_x, a_y, a_z)$ is some vector position operator. This quantity is represented as a 2×2 matrix:

$$\boldsymbol{\sigma} \cdot \mathbf{a} \equiv \sum_k a_k\, \sigma_k = \begin{pmatrix} +a_3 & a_1 - \mathrm{i}\, a_2 \\ a_1 + \mathrm{i}\, a_2 & -a_3 \end{pmatrix}. \tag{5.91}$$

Since, in the Schrödinger representation, a general position operator takes the form of a differential operator in x', y', or z', it is clear that the above quantity must be regarded as a matrix differential operator that acts on spinors of the general form (5.83). The important identity

$$(\boldsymbol{\sigma} \cdot \mathbf{a})\, (\boldsymbol{\sigma} \cdot \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} + \mathrm{i}\, \boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}) \tag{5.92}$$

follows from the commutation and anti-commutation relations (5.75) and (5.76). Thus,

$$\begin{aligned} \sum_j \sigma_j\, a_j \sum_k \sigma_k\, b_k &= \sum_j \sum_k \left(\frac{1}{2}\, \{\sigma_j, \sigma_k\} + \frac{1}{2} [\sigma_j, \sigma_k] \right) a_j\, b_k \\ &= \sum_j \sum_k (\delta_{jk} + \mathrm{i}\, \epsilon_{jkl}\, \sigma_l)\, a_j\, b_k \\ &= \mathbf{a} \cdot \mathbf{b} + \mathrm{i}\, \boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}). \end{aligned} \tag{5.93}$$

A general rotation operator in spin space is written

$$T(\varDelta\phi) = \exp\left(-\mathrm{i}\,\mathbf{S}\cdot\mathbf{n}\,\varDelta\varphi/\hbar\right), \tag{5.94}$$

by analogy with Equation (5.24), where $\mathbf{n}$ is a unit vector pointing along the axis of rotation, and $\varDelta\varphi$ is the angle of rotation. Here, $\mathbf{n}$ can be regarded as a trivial position operator. The rotation operator is represented

$$\exp\left(-\mathrm{i}\,\mathbf{S}\cdot\mathbf{n}\,\varDelta\varphi/\hbar\right) \rightarrow \exp\left(-\mathrm{i}\,\boldsymbol{\sigma}\cdot\mathbf{n}\,\varDelta\varphi/2\right) \tag{5.95}$$

in the Pauli scheme. The term on the right-hand side of the above expression is the exponential of a matrix. This can easily be evaluated using the Taylor series for an exponential, plus the rules

$$(\boldsymbol{\sigma}\cdot\mathbf{n})^k = 1 \qquad \text{for } k \text{ even}, \tag{5.96}$$

$$(\boldsymbol{\sigma}\cdot\mathbf{n})^k = (\boldsymbol{\sigma}\cdot\mathbf{n}) \qquad \text{for } k \text{ odd}. \tag{5.97}$$

These rules follow trivially from the identity (5.92). Thus, we can write

$$\begin{aligned}\exp\left(-\mathrm{i}\,\boldsymbol{\sigma}\cdot\mathbf{n}\,\varDelta\varphi/2\right) &= \left[1 - \frac{(\boldsymbol{\sigma}\cdot\mathbf{n})^2}{2!}\left(\frac{\varDelta\varphi}{2}\right)^2 + \frac{(\boldsymbol{\sigma}\cdot\mathbf{n})^4}{4!}\left(\frac{\varDelta\varphi}{2}\right)^4 + \cdots\right] \\ &\quad - \mathrm{i}\left[(\boldsymbol{\sigma}\cdot\mathbf{n})\left(\frac{\varDelta\varphi}{2}\right) - \frac{(\boldsymbol{\sigma}\cdot\mathbf{n})^3}{3!}\left(\frac{\varDelta\varphi}{2}\right)^3 + \cdots\right] \\ &= \cos(\varDelta\varphi/2)\,\mathbf{1} - \mathrm{i}\,\sin(\varDelta\varphi/2)\,\boldsymbol{\sigma}\cdot\mathbf{n}.\end{aligned} \tag{5.98}$$

The explicit 2×2 form of this matrix is

$$\begin{pmatrix} \cos(\varDelta\varphi/2) - \mathrm{i}\,n_z\sin(\varDelta\varphi/2) & (-\mathrm{i}\,n_x - n_y)\sin(\varDelta\varphi/2) \\ (-\mathrm{i}\,n_x + n_y)\sin(\varDelta\varphi/2) & \cos(\varDelta\varphi/2) + \mathrm{i}\,n_z\sin(\varDelta\varphi/2) \end{pmatrix}. \tag{5.99}$$

Rotation matrices act on spinors in much the same manner as the corresponding rotation operators act on state kets. Thus,

$$\chi' = \exp\left(-\mathrm{i}\,\boldsymbol{\sigma}\cdot\mathbf{n}\,\varDelta\varphi/2\right)\chi, \tag{5.100}$$

where χ' denotes the spinor obtained after rotating the spinor χ an angle $\varDelta\varphi$ about the **n**-axis. The Pauli matrices remain unchanged under rotations. However, the quantity $\chi^\dagger\,\sigma_k\,\chi$ is proportional to the expectation value of S_k [see Equation (5.74)], so we would expect it to transform like a vector under rotation (see Section 5.4). In fact, we require

$$(\chi^\dagger\,\sigma_k\,\chi)' \equiv (\chi^\dagger)'\sigma_k\,\chi' = \sum_l R_{kl}\,(\chi^\dagger\sigma_l\,\chi), \tag{5.101}$$

where the R_{kl} are the elements of a conventional rotation matrix. This is easily demonstrated, because

$$\exp\left(\frac{\mathrm{i}\,\sigma_3\,\varDelta\varphi}{2}\right)\sigma_1\exp\left(\frac{-\mathrm{i}\,\sigma_3\,\varDelta\varphi}{2}\right) = \sigma_1\cos\varDelta\varphi - \sigma_2\sin\varDelta\varphi \tag{5.102}$$

plus all cyclic permutations. The above expression is the 2×2 matrix analogue of (see Section 5.4)

$$\exp\left(\frac{\mathrm{i}\,S_z\,\Delta\varphi}{\hbar}\right)S_x\exp\left(\frac{-\mathrm{i}\,S_z\,\Delta\varphi}{\hbar}\right) = S_x\,\cos\Delta\varphi - S_y\,\sin\Delta\varphi. \tag{5.103}$$

The previous two formulae can both be validated using the Baker-Hausdorff lemma, (5.31), which holds for Hermitian matrices, in addition to Hermitian operators.

5.8 Spin Greater Than One-Half Systems

In the absence of spin, the Hamiltonian can be written as some function of the position and momentum operators. Using the Schrödinger representation, in which $\mathbf{p}\rightarrow -\mathrm{i}\,\hbar\,\nabla$, the energy eigenvalue problem,

$$H\,|E\rangle = E\,|E\rangle, \tag{5.104}$$

can be transformed into a partial differential equation for the wavefunction $\psi(\mathbf{x}')\equiv\langle\mathbf{x}'|E\rangle$. This function specifies the probability density for observing the particle at a given position, $\mathbf{x}'$. In general, we find

$$H\,\psi = E\,\psi, \tag{5.105}$$

where H is now a partial differential operator. The boundary conditions (for a bound state) are obtained from the normalization constraint

$$\int d^3x'\,|\psi|^2 = 1. \tag{5.106}$$

This is all very familiar. However, we now know how to generalize this scheme to deal with a spin one-half particle. Instead of representing the state of the particle by a single wavefunction, we use *two* wavefunctions. The first, $\psi_+(\mathbf{x}')$, specifies the probability density of observing the particle at position $\mathbf{x}'$ with spin angular momentum $+\hbar/2$ in the z-direction. The second, $\psi_-(\mathbf{x}')$, specifies the probability density of observing the particle at position $\mathbf{x}'$ with spin angular momentum $-\hbar/2$ in the z-direction. In the Pauli scheme, these wavefunctions are combined into a *spinor*, χ, which is simply the column vector of ψ_+ and ψ_-. In general, the Hamiltonian is a function of the position, momentum, and spin operators. Adopting the Schrödinger representation, and the Pauli scheme, the energy eigenvalue problem reduces to

$$H\,\chi = E\,\chi, \tag{5.107}$$

where χ is a spinor (i.e., a 2×1 matrix of wavefunctions) and H is a 2×2 matrix partial differential operator [see Equation (5.91)]. The above spinor equation can always be written out explicitly as *two coupled partial differential equations* for ψ_+ and ψ_-.

Suppose that the Hamiltonian has no dependence on the spin operators. In this case, the Hamiltonian is represented as *diagonal* 2×2 matrix partial differential operator in the Schrödinger/Pauli scheme [see Equation (5.90)]. In other words, the partial differential equation for ψ_+ decouples from that for ψ_-. In fact, both equations have the same form, so there is only really one differential equation. In this situation, the most general solution to Equation (5.107) can be written

$$\chi = \psi(\mathbf{x}') \begin{pmatrix} c_+ \\ c_- \end{pmatrix}. \tag{5.108}$$

Here, $\psi(\mathbf{x}')$ is determined by the solution of the differential equation, and the $c_\pm$ are arbitrary complex numbers. The physical significance of the above expression is clear. The Hamiltonian determines the relative probabilities of finding the particle at various different positions, but the direction of its spin angular momentum remains undetermined.

Suppose that the Hamiltonian depends only on the spin operators. In this case, the Hamiltonian is represented as a 2×2 matrix of complex numbers in the Schrödinger/Pauli scheme [see Equation (5.73)], and the spinor eigenvalue equation (5.107) reduces to a straightforward matrix eigenvalue problem. The most general solution can again be written

$$\chi = \psi(\mathbf{x}') \begin{pmatrix} c_+ \\ c_- \end{pmatrix}. \tag{5.109}$$

Here, the ratio c_+/c_- is determined by the matrix eigenvalue problem, and the wavefunction $\psi(\mathbf{x}')$ is arbitrary. Clearly, the Hamiltonian determines the direction of the particle's spin angular momentum, but leaves its position undetermined.

In general, of course, the Hamiltonian is a function of both position and spin operators. In this case, it is not possible to decompose the spinor as in Equations (5.108) and (5.109). In other words, a general Hamiltonian causes the direction of the particle's spin angular momentum to vary with position in some specified manner. This can only be represented as a spinor involving different wavefunctions, ψ_+ and ψ_-.

But, what happens if we have a spin one or a spin three-halves particle? It turns out that we can generalize the Pauli two-component scheme in a fairly straightforward manner. Consider a spin-s particle: i.e., a particle for which the eigenvalue of S^2 is $s\,(s+1)\,\hbar^2$. Here, s is either an integer, or a half-integer. The eigenvalues of S_z are written $s_z\,\hbar$, where s_z is allowed to take the values $s, s-1, \cdots, -s+1, -s$. In fact, there are $2\,s+1$ distinct allowed values of s_z. Not surprisingly, we can represent the state of the particle by $2\,s+1$ different wavefunctions, denoted $\psi_{s_z}(\mathbf{x}')$. Here, $\psi_{s_z}(\mathbf{x}')$ specifies the probability density for observing the particle at position $\mathbf{x}'$ with spin angular momentum $s_z\,\hbar$ in the z-direction. More exactly,

$$\psi_{s_z}(\mathbf{x}') = \langle \mathbf{x}' | \langle s, s_z || A \rangle \rangle, \tag{5.110}$$

where $||A\rangle\rangle$ denotes a state ket in the product space of the position and spin operators. The state of the particle can be represented more succinctly by a spinor, χ, which is simply the $2\,s+1$ component column vector of the $\psi_{s_z}(\mathbf{x}')$. Thus, a spin one-half particle is represented by a two-component spinor, a spin one particle by a three-component spinor, a spin three-halves particle by a four-component spinor, and so on.

In this extended Schrödinger/Pauli scheme, position space operators take the form of *diagonal* $(2\,s+1) \times (2\,s+1)$ matrix differential operators. Thus, we can represent the momentum operators as [see Equation (5.90)]

$$p_k \to -\mathrm{i}\,\hbar\,\frac{\partial}{\partial x_k'}\,\mathbf{1}, \tag{5.111}$$

where **1** is the $(2\,s+1)\times(2\,s+1)$ unit matrix. We represent the spin operators as

$$S_k \to s\,\hbar\,\sigma_k, \tag{5.112}$$

where the $(2\,s+1)\times(2\,s+1)$ extended Pauli matrix σ_k has elements

$$(\sigma_k)_{j\,l} = \frac{\langle s, j|\,S_k\,|s, l\rangle}{s\,\hbar}. \tag{5.113}$$

Here, j, l are integers, or half-integers, lying in the range $-s$ to $+s$. But, how can we evaluate the brackets $\langle s, j|\,S_k\,|s, l\rangle$ and, thereby, construct the extended Pauli matrices? In fact, it is trivial to construct the σ_z matrix. By definition,

$$S_z\,|s, j\rangle = j\,\hbar\,|s, j\rangle. \tag{5.114}$$

Hence,

$$(\sigma_3)_{j\,l} = \frac{\langle s, j|\,S_z\,|s, l\rangle}{s\,\hbar} = \frac{j}{s}\,\delta_{j\,l}, \tag{5.115}$$

where use has been made of the orthonormality property of the $|s, j\rangle$. Thus, σ_z is the suitably normalized diagonal matrix of the eigenvalues of S_z. The matrix elements of σ_x and σ_y are most easily obtained by considering the shift operators,

$$S^{\pm} = S_x \pm \mathrm{i}\,S_y. \tag{5.116}$$

We know, from Equations (4.55)–(4.56), that

$$S^{+}\,|s, j\rangle = [s\,(s+1) - j\,(j+1)]^{1/2}\,\hbar\,|s, j+1\rangle, \tag{5.117}$$

$$S^{-}\,|s, j\rangle = [s\,(s+1) - j\,(j-1)]^{1/2}\,\hbar\,|s, j-1\rangle. \tag{5.118}$$

It follows from Equations (5.113), and (5.116)–(5.118), that

$$(\sigma_1)_{j\,l} = \frac{[s\,(s+1) - j\,(j-1)]^{1/2}}{2\,s}\,\delta_{j\,l+1} + \frac{[s\,(s+1) - j\,(j+1)]^{1/2}}{2\,s}\,\delta_{j\,l-1}, \tag{5.119}$$

$$(\sigma_2)_{j\,l} = \frac{[s\,(s+1) - j\,(j-1)]^{1/2}}{2\,\mathrm{i}\,s}\,\delta_{j\,l+1} - \frac{[s\,(s+1) - j\,(j+1)]^{1/2}}{2\,\mathrm{i}\,s}\,\delta_{j\,l-1}. \tag{5.120}$$

According to Equations (5.115) and (5.119)–(5.120), the Pauli matrices for a spin one-half ($s = 1/2$) particle are

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{5.121}$$

$$\sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \tag{5.122}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{5.123}$$

as we have seen previously. For a spin one ($s = 1$) particle, we find that

$$\sigma_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{5.124}$$

$$\sigma_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix}, \tag{5.125}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{5.126}$$

In fact, we can now construct the Pauli matrices for a spin anything particle. This means that we can convert the general energy eigenvalue problem for a spin-s particle, where the Hamiltonian is some function of position and spin operators, into $2\,s+1$ coupled partial differential equations involving the $2\,s+1$ wavefunctions $\psi_{s_z}(\mathbf{x}')$. Unfortunately, such a system of equations is generally too complicated to solve exactly.

Exercises

5.1 Demonstrate that the operators defined in Equations (5.11)–(5.13) are Hermitian, and satisfy the commutation relations (5.1).

5.2 Prove the Baker-Hausdorff lemma, (5.31).

5.3 Find the Pauli representations of the normalized eigenstates of S_x and S_y for a spin-1/2 particle.

5.4 Suppose that a spin-1/2 particle has a spin vector that lies in the x-z plane, making an angle θ with the z-axis. Demonstrate that a measurement of S_z yields $\hbar/2$ with probability $\cos^2(\theta/2)$, and $-\hbar/2$ with probability $\sin^2(\theta/2)$.

5.5 An electron is in the spin-state

$$\chi = A\begin{pmatrix} 1-2\,\mathrm{i} \\ 2 \end{pmatrix}$$

in the Pauli representation. Determine the constant A by normalizing χ. If a measurement of S_z is made, what values will be obtained, and with what probabilities? What is the expectation value of S_z? Repeat the above calculations for S_x and S_y.

5.6 Consider a spin-1/2 system represented by the normalized spinor

$$\chi = \begin{pmatrix} \cos\alpha \\ \sin\alpha\,\exp(\,\mathrm{i}\beta) \end{pmatrix}$$

in the Pauli representation, where α and β are real. What is the probability that a measurement of S_y yields $-\hbar/2$?

5.7 An electron is at rest in an oscillating magnetic field

$$\mathbf{B} = B_0 \cos(\omega t)\,\mathbf{e}_z,$$

where B_0 and ω are real positive constants.

(a) Find the Hamiltonian of the system.

(b) If the electron starts in the spin-up state with respect to the x-axis, determine the spinor $\chi(t)$ that represents the state of the system in the Pauli representation at all subsequent times.

(c) Find the probability that a measurement of S_x yields the result $-\hbar/2$ as a function of time.

(d) What is the minimum value of B_0 required to force a complete flip in S_x?

6 Addition of Angular Momentum

6.1 Introduction

Consider a hydrogen atom whose orbiting electron is in an $l = 1$ state. The electron, consequently, possesses orbital angular momentum of magnitude $\hbar$, and spin angular momentum of magnitude $\hbar/2$. So, what is the electron's total angular momentum?

6.2 Analysis

In order to answer this question, we need to learn how to add angular momentum operators. Consider the most general case. Suppose that we have two sets of angular momentum operators, $\mathbf{J}_1$ and $\mathbf{J}_2$. By definition, these operators are Hermitian, and obey the fundamental commutation relations

$$\mathbf{J}_1 \times \mathbf{J}_1 = \mathrm{i}\,\hbar\,\mathbf{J}_1, \tag{6.1}$$

$$\mathbf{J}_2 \times \mathbf{J}_2 = \mathrm{i}\,\hbar\,\mathbf{J}_2. \tag{6.2}$$

Let us assume that the two groups of operators correspond to different degrees of freedom of the system, so that

$$[J_{1\,i}, J_{2\,j}] = 0, \tag{6.3}$$

where i, j stand for either x, y, or z. For instance, $\mathbf{J}_1$ could be an orbital angular momentum operator, and $\mathbf{J}_2$ a spin angular momentum operator. Alternatively, $\mathbf{J}_1$ and $\mathbf{J}_2$ could be the orbital angular momentum operators of two different particles in a multi-particle system. We know, from the general properties of angular momentum, that the eigenvalues of $J_1^{\,2}$ and $J_2^{\,2}$ can be written $j_1\,(j_1+1)\,\hbar^2$ and $j_2\,(j_2+1)\,\hbar^2$, respectively, where j_1 and j_2 are either integers, or half-integers. We also know that the eigenvalues of $J_{1\,z}$ and $J_{2\,z}$ take the form $m_1\,\hbar$ and $m_2\,\hbar$, respectively, where m_1 and m_2 are numbers lying in the ranges $j_1, j_1 - 1, \cdots, -j_1 + 1, -j_1$ and $j_2, j_2 - 1, \cdots, -j_2 + 1, -j_2$, respectively.

Let us define the total angular momentum operator

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2. \tag{6.4}$$

Now, $\mathbf{J}$ is an Hermitian operator, because it is the sum of Hermitian operators. Moreover, according to Equations (4.11) and (4.14), $\mathbf{J}$ satisfies the fundamental commutation relation

$$\mathbf{J} \times \mathbf{J} = \mathrm{i}\,\hbar\,\mathbf{J}. \tag{6.5}$$

Thus, $\mathbf{J}$ possesses all of the expected properties of an angular momentum operator. It follows that the eigenvalue of $J^{\,2}$ can be written $j\,(j+1)\,\hbar^2$, where j is an integer, or a half-integer. Moreover, the eigenvalue of J_z takes the form $m\,\hbar$, where m lies in the range $j, j-1, \cdots, -j+1, -j$. At this stage, however, we do not know the relationship between the quantum numbers of the total angular momentum, j and m, and those of the individual angular momenta, j_1, j_2, m_1, and m_2.

Now,

$$J^2 = J_1^2 + J_2^2 + 2\,\mathbf{J}_1\cdot\mathbf{J}_2. \tag{6.6}$$

Furthermore, we know that

$$[J_1^2, J_{1\,i}] = 0, \tag{6.7}$$

$$[J_2^2, J_{2\,i}] = 0, \tag{6.8}$$

and also that all of the $J_{1\,i}$, J_1^2 operators commute with the $J_{2\,i}$, J_2^2 operators. It follows from Equation (6.6) that

$$[J^2, J_1^2] = [J^2, J_2^2] = 0. \tag{6.9}$$

This implies that the quantum numbers j_1, j_2, and j can all be measured simultaneously. In other words, we can know the magnitude of the total angular momentum together with the magnitudes of the component angular momenta. However, it is apparent from Equation (6.6) that

$$[J^2, J_{1\,z}] \neq 0, \tag{6.10}$$

$$[J^2, J_{2\,z}] \neq 0. \tag{6.11}$$

This suggests that it is not possible to measure the quantum numbers m_1 and m_2 simultaneously with the quantum number j. Thus, we cannot determine the projections of the individual angular momenta along the z-axis at the same time as the magnitude of the total angular momentum.

It is clear, from the preceding discussion, that we can form two alternate groups of mutually commuting operators. The first group is $J_1^2, J_2^2, J_{1\,z}$, and $J_{2\,z}$. The second group is J_1^2, J_2^2, J^2, and J_z. These two groups of operators are incompatible with one another. We can define simultaneous eigenkets of each operator group. The simultaneous eigenkets of J_1^2, J_2^2, J_{1z}, and J_{2z} are denoted $|j_1, j_2; m_1, m_2\rangle$, where

$$J_1^2\,|j_1, j_2; m_1, m_2\rangle = j_1\,(j_1+1)\,\hbar^2\,|j_1, j_2; m_1, m_2\rangle, \tag{6.12}$$

$$J_2^2\,|j_1, j_2; m_1, m_2\rangle = j_2\,(j_2+1)\,\hbar^2\,|j_1, j_2; m_1, m_2\rangle, \tag{6.13}$$

$$J_{1z}\,|j_1, j_2; m_1, m_2\rangle = m_1\,\hbar\,|j_1, j_2; m_1, m_2\rangle, \tag{6.14}$$

$$J_{2z}\,|j_1, j_2; m_1, m_2\rangle = m_2\,\hbar\,|j_1, j_2; m_1, m_2\rangle. \tag{6.15}$$

The simultaneous eigenkets of J_1^2, J_2^2, J^2 and J_z are denoted $|j_1, j_2; j, m\rangle$, where

$$J_1^2\,|j_1, j_2; j, m\rangle = j_1\,(j_1+1)\,\hbar^2\,|j_1, j_2; j, m\rangle, \tag{6.16}$$

$$J_2^2\,|j_1, j_2; j, m\rangle = j_2\,(j_2+1)\,\hbar^2\,|j_1, j_2; j, m\rangle, \tag{6.17}$$

$$J^2\,|j_1, j_2; j, m\rangle = j\,(j+1)\,\hbar^2\,|j_1, j_2; j, m\rangle, \tag{6.18}$$

$$J_z\,|j_1, j_2; j, m\rangle = m\,\hbar\,|j_1, j_2; j, m\rangle. \tag{6.19}$$

Each set of eigenkets are complete, mutually orthogonal (for eigenkets corresponding to different sets of eigenvalues), and have unit norms. Since the operators J_1^2 and J_2^2 are common to both operator groups, we can assume that the quantum numbers j_1 and j_2 are known. In other words, we can always determine the magnitudes of the individual angular momenta. In addition, we can either know the quantum numbers m_1 and m_2, or the quantum numbers j and m, but we cannot know both pairs of quantum numbers at the same time. We can write a conventional completeness relation for both sets of eigenkets:

$$\sum_{m_1}\sum_{m_2} |j_1, j_2; m_1, m_2\rangle\langle j_1, j_2; m_1, m_2| = 1, \tag{6.20}$$

$$\sum_{j}\sum_{m} |j_1, j_2; j, m\rangle\langle j_1, j_2; j, m| = 1, \tag{6.21}$$

where the right-hand sides denote the identity operator in the ket space corresponding to states of given j_1 and j_2. The summation is over all allowed values of m_1, m_2, j, and m.

As we have seen, the operator group $J_1^{\,2}$, $J_2^{\,2}$, $J^{\,2}$, and J_z is incompatible with the group $J_1^{\,2}$, $J_2^{\,2}$, $J_{1\,z}$, and $J_{2\,z}$. This means that if the system is in a simultaneous eigenstate of the former group then, in general, it is not in an eigenstate of the latter. In other words, if the quantum numbers j_1, j_2, j, and m are known with certainty then a measurement of the quantum numbers m_1 and m_2 will give a range of possible values. We can use the completeness relation (6.20) to write

$$|j_1, j_2; j, m\rangle = \sum_{m_1}\sum_{m_2}\langle j_1, j_2; m_1, m_2|j_1, j_2; j, m\rangle|j_1, j_2; m_1, m_2\rangle. \tag{6.22}$$

Thus, we can write the eigenkets of the first group of operators as a weighted sum of the eigenkets of the second set. The weights, $\langle j_1, j_2; m_1, m_2|j_1, j_2; j, m\rangle$, are called the *Clebsch-Gordon coefficients*. If the system is in a state where a measurement of $J_1^{\,2}, J_2^{\,2}, J^{\,2}$, and J_z is bound to give the results $j_1\,(j_1+1)\,\hbar^2$, $j_2\,(j_2+1)\,\hbar^2$, $j\,(j+1)\,\hbar^2$, and $j_z\,\hbar$, respectively, then a measurement of $J_{1\,z}$ and $J_{2\,z}$ will give the results $m_1\,\hbar$ and $m_2\,\hbar$, respectively, with probability $|\langle j_1, j_2; m_1, m_2|j_1, j_2; j, m\rangle|^2$.

The Clebsch-Gordon coefficients possess a number of very important properties. First, the coefficients are zero unless

$$m = m_1 + m_2. \tag{6.23}$$

To prove this, we note that

$$(J_z - J_{1\,z} - J_{2\,z})\,|j_1, j_2; j, m\rangle = 0. \tag{6.24}$$

Forming the inner product with $\langle j_1, j_2; m_1, m_2|$, we obtain

$$(m - m_1 - m_2)\,\langle j_1, j_2; m_1, m_2|j_1, j_2; j, m\rangle = 0, \tag{6.25}$$

which proves the assertion. Thus, the z-components of different angular momenta add algebraically. So, an electron in an $l = 1$ state, with orbital angular momentum $\hbar$, and spin angular momentum $\hbar/2$, projected along the z-axis, constitutes a state whose total angular momentum projected along the z-axis is $3\,\hbar/2$. What is uncertain is the magnitude of the total angular momentum.

Second, the coefficients vanish unless

$$|j_1 - j_2| \leq j \leq j_1 + j_2. \tag{6.26}$$

We can assume, without loss of generality, that $j_1 \geq j_2$. We know, from Equation (6.23), that for given j_1 and j_2 the largest possible value of m is $j_1 + j_2$ (because j_1 is the largest possible value of m_1, etc.). This implies that the largest possible value of j is $j_1 + j_2$ (since, by definition, the largest value of m is equal to j). Now, there are $(2\,j_1 + 1)$ allowable values of m_1 and $(2\,j_2 + 1)$ allowable values of m_2. Thus, there are $(2\,j_1 + 1)\,(2\,j_2 + 1)$ independent eigenkets, $|j_1, j_2; m_1, m_2\rangle$, needed to span the ket space corresponding to fixed j_1 and j_2. Because the eigenkets $|j_1, j_2; j, m\rangle$ span the same space, they must also form a set of $(2\,j_1 + 1)\,(2\,j_2 + 1)$ independent kets. In other words, there can only be $(2\,j_1 + 1)\,(2\,j_2 + 1)$ distinct allowable values of the quantum numbers j and m. For each allowed value of j, there are $2\,j + 1$ allowed values of m. We have already seen that the maximum allowed value of j is $j_1 + j_2$. It is easily seen that if the minimum allowed value of j is $j_1 - j_2$ then the total number of allowed values of j and m is $(2\,j_1 + 1)\,(2\,j_2 + 1)$: i.e.,

$$\sum_{j=j_1-j_2,j_1+j_2} (2\,j + 1) \equiv (2\,j_1 + 1)\,(2\,j_2 + 1). \tag{6.27}$$

This proves our assertion.

Third, the sum of the modulus squared of all of the Clebsch-Gordon coefficients is unity: i.e.,

$$\sum_{m_1}\sum_{m_2} |\langle j_1, j_2; m_1, m_2|j_1, j_2; j, m\rangle|^2 = 1. \tag{6.28}$$

This assertion is proved as follows:

$$\begin{aligned}\langle j_1, j_2; j, m|j_1, j_2; j, m\rangle &= \sum_{m_1}\sum_{m_2}\langle j_1, j_2; j, m|j_1, j_2; m_1, m_2\rangle\langle j_1, j_2; m_1, m_2|j_1, j_2; j, m\rangle \\ &= \sum_{m_1}\sum_{m_2} |\langle j_1, j_2; m_1, m_2|j_1, j_2; j, m\rangle|^2 = 1,\end{aligned} \tag{6.29}$$

where use has been made of the completeness relation (6.20).

Finally, the Clebsch-Gordon coefficients obey two recursion relations. To obtain these relations, we start from

$$J^{\pm}|j_1, j_2; j, m\rangle = (J_1^{\pm} + J_2^{\pm})\sum_{m_1'}\sum_{m_2'}\langle j_1, j_2; m_1', m_2'|j_1, j_2; j, m\rangle|j_1, j_2; m_1', m_2'\rangle. \tag{6.30}$$

Making use of the well-known properties of the shift operators, which are specified by Equations (4.55)–(4.56), we obtain

$$\begin{aligned}\sqrt{j\,(j+1) - m\,(m \pm 1)}\,|j_1, j_2; j, m \pm 1\rangle = \sum_{m_1'}\sum_{m_2'}\Big(&\sqrt{j_1\,(j_1+1) - m_1'\,(m_1' \pm 1)}\,|j_1, j_2; m_1' \pm 1, m_2'\rangle \\ &+ \sqrt{j_2\,(j_2+1) - m_2'\,(m_2' \pm 1)}\,|j_1, j_2; m_1', m_2' \pm 1\rangle\Big) \\ &\times \langle j_1, j_2; m_1', m_2'|j_1, j_2; j, m\rangle.\end{aligned} \tag{6.31}$$

Taking the inner product with $\langle j_1, j_2; m_1, m_2|$, and making use of the orthonormality property of the basis eigenkets, we obtain the desired recursion relations:

$$\begin{aligned}\sqrt{j\,(j+1) - m\,(m \pm 1)}\,\langle j_1, j_2; m_1, m_2|j_1, j_2; j, m \pm 1\rangle = &\sqrt{j_1\,(j_1+1) - m_1\,(m_1 \mp 1)}\\ &\times \langle j_1, j_2; m_1 \mp 1, m_2|j_1, j_2; j, m\rangle\\ &+ \sqrt{j_2\,(j_2+1) - m_2\,(m_2 \mp 1)}\\ &\times \langle j_1, j_2; m_1, m_2 \mp 1|j_1, j_2; j, m\rangle. \qquad (6.32)\end{aligned}$$

It is clear, from the absence of complex coupling coefficients in the above relations, that we can always choose the Clebsch-Gordon coefficients to be real numbers. This is convenient, because it ensures that the inverse Clebsch-Gordon coefficients, $\langle j_1, j_2; j, m|j_1, j_2; m_1, m_2\rangle$, are identical to the Clebsch-Gordon coefficients. In other words,

$$\langle j_1, j_2; j, m|j_1, j_2; m_1, m_2\rangle = \langle j_1, j_2; m_1, m_2|j_1, j_2; j, m\rangle. \qquad (6.33)$$

The inverse Clebsch-Gordon coefficients are the weights in the expansion of the $|j_1, j_2; m_1, m_2\rangle$ in terms of the $|j_1, j_2; j, m\rangle$:

$$|j_1, j_2; m_1, m_2\rangle = \sum_j \sum_m \langle j_1, j_2; j, m|j_1, j_2; m_1, m_2\rangle |j_1, j_2; j, m\rangle. \qquad (6.34)$$

It turns out that the recursion relations (6.32), together with the normalization condition (6.28), are sufficient to completely determine the Clebsch-Gordon coefficients to within an arbitrary sign (multiplied into all of the coefficients). This sign is fixed by convention. The easiest way of demonstrating this assertion is by considering a specific example.

Let us add the angular momentum of two spin one-half systems: e.g., two electrons at rest. So, $j_1 = j_2 = 1/2$. We know, from general principles, that $|m_1| \leq 1/2$ and $|m_2| \leq 1/2$. We also know, from Equation (6.26), that $0 \leq j \leq 1$, where the allowed values of j differ by integer amounts. It follows that either $j = 0$ or $j = 1$. Thus, two spin one-half systems can be combined to form either a spin zero system or a spin one system. It is helpful to arrange all of the possibly non-zero Clebsch-Gordon coefficients in a table:

m_1	m_2				
1/2	1/2	?	?	?	?
1/2	-1/2	?	?	?	?
-1/2	1/2	?	?	?	?
-1/2	-1/2	?	?	?	?
j_1=1/2	j	1	1	1	0
j_2=1/2	m	1	0	-1	0

The box in this table corresponding to $m_1 = 1/2, m_2 = 1/2, j = 1, m = 1$ gives the Clebsch-Gordon coefficient $\langle 1/2, 1/2; 1/2, 1/2|1/2, 1/2; 1, 1\rangle$, or the inverse Clebsch-Gordon coefficient $\langle 1/2, 1/2; 1, 1|1/2, 1/2; 1/2, 1/2\rangle$. All the boxes contain question marks because, at this stage, we do not know the values of any Clebsch-Gordon coefficients.

A Clebsch-Gordon coefficient is automatically zero unless $m_1 + m_2 = m$. In other words, the z-components of angular momentum have to add algebraically. Many of the boxes in the above table correspond to $m_1 + m_2 \neq m$. We immediately conclude that these boxes must contain zeroes: i.e.,

m_1	m_2				
1/2	1/2	?	0	0	0
1/2	-1/2	0	?	0	?
-1/2	1/2	0	?	0	?
-1/2	-1/2	0	0	?	0
j_1=1/2	j	1	1	1	0
j_2=1/2	m	1	0	-1	0

The normalization condition (6.28) implies that the sum of the squares of all the rows and columns of the above table must be unity. There are two rows and two columns that only contain a single non-zero entry. We conclude that these entries must be ±1, but we have no way of determining the signs at present. Thus,

m_1	m_2				
1/2	1/2	±1	0	0	0
1/2	-1/2	0	?	0	?
-1/2	1/2	0	?	0	?
-1/2	-1/2	0	0	±1	0
j_1=1/2	j	1	1	1	0
j_2=1/2	m	1	0	-1	0

Let us evaluate the recursion relation (6.32) for $j_1 = j_2 = 1/2$, with $j = 1$, $m = 0$, $m_1 = m_2 = \pm 1/2$, taking the upper/lower sign. We find that

$$\langle 1/2, -1/2|1, 0\rangle + \langle -1/2, 1/2|1, 0\rangle = \sqrt{2}\,\langle 1/2, 1/2|1, 1\rangle = \pm\sqrt{2}, \tag{6.35}$$

and

$$\langle 1/2, -1/2|1, 0\rangle + \langle -1/2, 1/2|1, 0\rangle = \sqrt{2}\,\langle -1/2, -1/2|1, -1\rangle = \pm\sqrt{2}. \tag{6.36}$$

Here, the j_1 and j_2 labels have been suppressed for ease of notation. We also know that

$$\langle 1/2, -1/2|1, 0\rangle^2 + \langle -1/2, 1/2|1, 0\rangle^2 = 1, \tag{6.37}$$

from the normalization condition. The only real solutions to the above set of equations are

$$\begin{aligned}\sqrt{2}\,\langle 1/2, -1/2|1, 0\rangle &= \sqrt{2}\,\langle -1/2, 1/2|1, 0\rangle \\ &= \langle 1/2, 1/2|1, 1\rangle = \langle 1/2, 1/2|1, -1\rangle = \pm 1.\end{aligned} \tag{6.38}$$

The choice of sign is arbitrary—the conventional choice is a positive sign. Thus, our table now reads

m_1	m_2				
1/2	1/2	1	0	0	0
1/2	-1/2	0	$1/\sqrt{2}$	0	?
-1/2	1/2	0	$1/\sqrt{2}$	0	?
-1/2	-1/2	0	0	1	0
j_1=1/2	j	1	1	1	0
j_2=1/2	m	1	0	-1	0

We could fill in the remaining unknown entries of our table by using the recursion relation again. However, an easier method is to observe that the rows and columns of the table must all be mutually orthogonal. That is, the dot product of a row with any other row must be zero. Likewise, for the dot product of a column with any other column. This follows because the entries in the table give the expansion coefficients of one of our alternative sets of eigenkets in terms of the other set, and each set of eigenkets contains mutually orthogonal vectors with unit norms. The normalization condition tells us that the dot product of a row or column with itself must be unity. The only way that the dot product of the fourth column with the second column can be zero is if the unknown entries are equal and opposite. The requirement that the dot product of the fourth column with itself is unity tells us that the magnitudes of the unknown entries have to be $1/\sqrt{2}$. The unknown entries are undetermined to an arbitrary sign multiplied into them both. Thus, the final form of our table (with the conventional choice of arbitrary signs) is

m_1	m_2				
1/2	1/2	1	0	0	0
1/2	-1/2	0	$1/\sqrt{2}$	0	$1/\sqrt{2}$
-1/2	1/2	0	$1/\sqrt{2}$	0	$-1/\sqrt{2}$
-1/2	-1/2	0	0	1	0
j_1=1/2	j	1	1	1	0
j_2=1/2	m	1	0	-1	0

The table can be read in one of two ways. The columns give the expansions of the eigenstates of overall angular momentum in terms of the eigenstates of the individual angular momenta of the two component systems. Thus, the second column tells us that

$$|1,0\rangle = \frac{1}{\sqrt{2}}\left(|1/2,-1/2\rangle + |-1/2,1/2\rangle\right). \tag{6.39}$$

The ket on the left-hand side is a $|j,m\rangle$ ket, whereas those on the right-hand side are $|m_1,m_2\rangle$ kets. The rows give the expansions of the eigenstates of individual angular momentum in terms of those of overall angular momentum. Thus, the second row tells us that

$$|1/2,-1/2\rangle = \frac{1}{\sqrt{2}}\left(|1,0\rangle + |0,0\rangle\right). \tag{6.40}$$

Here, the ket on the left-hand side is a $|m_1, m_2\rangle$ ket, whereas those on the right-hand side are $|j, m\rangle$ kets.

Note that our table is really a combination of two sub-tables, one involving $j = 0$ states, and one involving $j = 1$ states. The Clebsch-Gordon coefficients corresponding to two different choices of j are completely independent: i.e., there is no recursion relation linking Clebsch-Gordon coefficients corresponding to different values of j. Thus, for every choice of j_1, j_2, and j we can construct a table of Clebsch-Gordon coefficients corresponding to the different allowed values of m_1, m_2, and m (subject to the constraint that $m_1 + m_2 = m$). A complete knowledge of angular momentum addition is equivalent to a knowing all possible tables of Clebsch-Gordon coefficients. These tables are listed (for moderate values of j_1, j_2 and j) in many standard reference books.

Exercises

6.1 Calculate the Clebsch-Gordon coefficients for adding spin one-half to spin one.

6.2 Calculate the Clebsch-Gordon coefficients for adding spin one to spin one.

6.3 An electron in a hydrogen atom occupies the combined spin and position state whose wavefunction is

$$\psi = R_{2\,1}(r)\left[\sqrt{1/3}\,Y_{1\,0}(\theta,\varphi)\,\chi_{+} + \sqrt{2/3}\,Y_{1\,1}(\theta,\varphi)\,\chi_{-}\right].$$

(a) What values would a measurement of L^2 yield, and with what probabilities?

(b) Same for L_z.

(c) Same for S^2.

(d) Same for S_z.

(e) Same for J^2.

(f) Same for J_z.

(g) What is the probability density for finding the electron at r, θ, φ?

(h) What is the probability density for finding the electron in the spin up state (with respect to the z-axis) at radius r?

6.4 In a low energy neutron-proton system (with zero orbital angular momentum) the potential energy is given by

$$V(\mathbf{x}) = V_1(r) + V_2(r)\left[3\,\frac{(\boldsymbol{\sigma}_n\cdot\mathbf{x})\,(\boldsymbol{\sigma}_p\cdot\mathbf{x})}{r^2} - \boldsymbol{\sigma}_n\cdot\boldsymbol{\sigma}_p\right] + V_3(r)\,\boldsymbol{\sigma}_n\cdot\boldsymbol{\sigma}_p,$$

where $r = |\mathbf{x}|$, $\boldsymbol{\sigma}_n$ denotes the vector of the Pauli matrices of the neutron, and $\boldsymbol{\sigma}_p$ denotes the vector of the Pauli matrices of the proton. Calculate the potential energy for the neutron-proton system:

(a) In the spin singlet (i.e., spin zero) state.

(b) In the spin triplet (i.e., spin one) state.

[Hint: Calculate the expectation value of $V(\mathbf{x})$ with respect to the overall spin state.]

6.5 Consider two electrons in a spin singlet (i.e., spin zero) state.

(a) If a measurement of the spin of one of the electrons shows that it is in the state with $S_z = \hbar/2$, what is the probability that a measurement of the z-component of the spin of the other electron yields $S_z = \hbar/2$?

(b) If a measurement of the spin of one of the electrons shows that it is in the state with $S_y = \hbar/2$, what is the probability that a measurement of the x-component of the spin of the other electron yields $S_x = -\hbar/2$?

(c) Finally, if electron 1 is in a spin state described by $\cos\alpha_1\,\chi_+ + \sin\alpha_1\,\mathrm{e}^{\,\mathrm{i}\beta_1}\,\chi_-$, and electron 2 is in a spin state described by $\cos\alpha_2\,\chi_+ + \sin\alpha_2\,\mathrm{e}^{\,\mathrm{i}\beta_2}\,\chi_-$, what is the probability that the two-electron spin state is a triplet (i.e., spin one) state?

7 Time-Independent Perturbation Theory

7.1 Introduction

We have developed techniques by which the general energy eigenvalue problem can be reduced to a set of coupled partial differential equations involving various wavefunctions. Unfortunately, the number of such problems that yield exactly soluble equations is comparatively small. It is, therefore, necessary to develop techniques for finding approximate solutions to otherwise intractable problems.

Consider the following problem, which is very common. The Hamiltonian of a system is written

$$H = H_0 + H_1. \tag{7.1}$$

Here, H_0 is a simple Hamiltonian for which we know the *exact* eigenvalues and eigenstates. H_1 introduces some interesting additional physics into the problem, but it is sufficiently complicated that when we add it to H_0 we can no longer find the exact energy eigenvalues and eigenstates. However, H_1 can, in some sense (which we shall specify more exactly later on), be regarded as being *small* compared to H_0. Let us try to find approximate eigenvalues and eigenstates of the modified Hamiltonian, $H_0 + H_1$, by performing a perturbation analysis about the eigenvalues and eigenstates of the original Hamiltonian, H_0.

7.2 Two-State System

Let us start by considering *time-independent perturbation theory*, in which the modification to the Hamiltonian, H_1, has no explicit dependence on time. It is usually assumed that the unperturbed Hamiltonian, H_0, is also time-independent.

Consider the simplest non-trivial system, in which there are only *two* independent eigenkets of the unperturbed Hamiltonian. These are denoted

$$H_0\,|1\rangle = E_1\,|1\rangle, \tag{7.2}$$

$$H_0\,|2\rangle = E_2\,|2\rangle. \tag{7.3}$$

It is assumed that these states, and their associated eigenvalues, are known. Because H_0 is, by definition, an Hermitian operator, its two eigenkets are mutually orthogonal and form a complete set. The lengths of these eigenkets are both normalized to unity. Let us now try to solve the modified energy eigenvalue problem

$$(H_0 + H_1)\,|E\rangle = E\,|E\rangle. \tag{7.4}$$

In fact, we can solve this problem exactly. Since the eigenkets of H_0 form a complete set, we can write

$$|E\rangle = \langle 1|E\rangle|1\rangle + \langle 2|E\rangle|2\rangle. \tag{7.5}$$

Right-multiplication of Equation (7.4) by $\langle 1|$ and $\langle 2|$ yields two coupled equations, which can be written in matrix form:

$$\begin{pmatrix} E_1 - E + e_{11} & e_{12} \\ e_{12}^{\ast} & E_2 - E + e_{22} \end{pmatrix} \begin{pmatrix} \langle 1|E\rangle \\ \langle 2|E\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{7.6}$$

Here,

$$e_{11} = \langle 1|\, H_1\, |1\rangle, \tag{7.7}$$

$$e_{22} = \langle 2|\, H_1\, |2\rangle, \tag{7.8}$$

$$e_{12} = \langle 1|\, H_1\, |2\rangle. \tag{7.9}$$

In the special (but common) case of a perturbing Hamiltonian whose diagonal matrix elements (in the unperturbed eigenstates) are zero, so that

$$e_{11} = e_{22} = 0, \tag{7.10}$$

the solution of Equation (7.6) (obtained by setting the determinant of the matrix equal to zero) is

$$E = \frac{(E_1 + E_2) \pm \sqrt{(E_1 - E_2)^2 + 4\,|e_{12}|^2}}{2}. \tag{7.11}$$

Let us expand in the supposedly small parameter

$$\epsilon = \frac{|e_{12}|}{|E_1 - E_2|}. \tag{7.12}$$

We obtain

$$E \simeq \frac{1}{2}\,(E_1 + E_2) \pm \frac{1}{2}\,(E_1 - E_2)\,(1 + 2\,\epsilon^2 + \cdots). \tag{7.13}$$

The above expression yields the modifications to the energy eigenvalues due to the perturbing Hamiltonian:

$$E_1' = E_1 + \frac{|e_{12}|^2}{E_1 - E_2} + \cdots\,, \tag{7.14}$$

$$E_2' = E_2 - \frac{|e_{12}|^2}{E_1 - E_2} + \cdots\,. \tag{7.15}$$

Note that H_1 causes the upper eigenvalue to rise, and the lower eigenvalue to fall. It is easily demonstrated that the modified eigenkets take the form

$$|1\rangle' = |1\rangle + \frac{e_{12}^{\ast}}{E_1 - E_2}\,|2\rangle + \cdots\,, \tag{7.16}$$

$$|2\rangle' = |2\rangle - \frac{e_{12}}{E_1 - E_2}\,|1\rangle + \cdots\,. \tag{7.17}$$

Thus, the modified energy eigenstates consist of one of the unperturbed eigenstates with a slight admixture of the other. Note that the series expansion in Equation (7.13) only converges if $2\,|\epsilon| < 1$. This suggests that the condition for the validity of the perturbation expansion is

$$|e_{12}| < \frac{|E_1 - E_2|}{2}. \tag{7.18}$$

In other words, when we say that H_1 needs to be small compared to H_0, what we really mean is that the above inequality needs to be satisfied.

7.3 Non-Degenerate Perturbation Theory

Let us now generalize our perturbation analysis to deal with systems possessing more than two energy eigenstates. The energy eigenstates of the unperturbed Hamiltonian, H_0, are denoted

$$H_0\,|n\rangle = E_n\,|n\rangle, \tag{7.19}$$

where n runs from 1 to N. The eigenkets $|n\rangle$ are orthogonal, form a complete set, and have their lengths normalized to unity. Let us now try to solve the energy eigenvalue problem for the perturbed Hamiltonian:

$$(H_0 + H_1)\,|E\rangle = E\,|E\rangle. \tag{7.20}$$

We can express $|E\rangle$ as a linear superposition of the unperturbed energy eigenkets,

$$|E\rangle = \sum_k \langle k|E\rangle |k\rangle, \tag{7.21}$$

where the summation is from $k = 1$ to N. Substituting the above equation into Equation (7.20), and right-multiplying by $\langle m|$, we obtain

$$(E_m + e_{mm} - E)\,\langle m|E\rangle + \sum_{k\neq m} e_{mk}\,\langle k|E\rangle = 0, \tag{7.22}$$

where

$$e_{mk} = \langle m|\,H_1\,|k\rangle. \tag{7.23}$$

Let us now develop our perturbation expansion. We assume that

$$\frac{|e_{mk}|}{E_m - E_k} \sim O(\epsilon), \tag{7.24}$$

for all $m \neq k$, where $\epsilon \ll 1$ is our expansion parameter. We also assume that

$$\frac{|e_{mm}|}{E_m} \sim O(\epsilon), \tag{7.25}$$

for all m. Let us search for a modified version of the nth unperturbed energy eigenstate, for which

$$E = E_n + O(\epsilon), \tag{7.26}$$

and

$$\langle n|E\rangle = 1, \tag{7.27}$$

$$\langle m|E\rangle \sim O(\epsilon), \tag{7.28}$$

for $m \neq n$. Suppose that we write out Equation (7.22) for $m \neq n$, neglecting terms that are $O(\epsilon^2)$ according to our expansion scheme. We find that

$$(E_m - E_n)\,\langle m|E\rangle + e_{mn} \simeq 0, \tag{7.29}$$

giving

$$\langle m|E\rangle \simeq -\frac{e_{mn}}{E_m - E_n}. \tag{7.30}$$

Substituting the above expression into Equation (7.22), evaluated for $m = n$, and neglecting $O(\epsilon^3)$ terms, we obtain

$$(E_n + e_{nn} - E) - \sum_{k\neq n} \frac{|e_{nk}|^2}{E_k - E_n} \simeq 0. \tag{7.31}$$

Thus, the modified nth energy eigenstate possesses an eigenvalue

$$E_n' = E_n + e_{nn} + \sum_{k\neq n} \frac{|e_{nk}|^2}{E_n - E_k} + O(\epsilon^3), \tag{7.32}$$

and a eigenket

$$|n\rangle' = |n\rangle + \sum_{k\neq n} \frac{e_{kn}}{E_n - E_k}\,|k\rangle + O(\epsilon^2). \tag{7.33}$$

Note that

$$\langle m|n\rangle' = \delta_{mn} + \frac{e_{nm}^*}{E_m - E_n} + \frac{e_{mn}}{E_n - E_m} + O(\epsilon^2) = \delta_{mn} + O(\epsilon^2). \tag{7.34}$$

Thus, the modified eigenkets remain orthogonal and properly normalized to $O(\epsilon^2)$.

7.4 Quadratic Stark Effect

Suppose that a one-electron atom [i.e., either a hydrogen atom, or an alkali metal atom (which possesses one valance electron orbiting outside a closed, spherically symmetric, shell)] is subjected to a uniform electric field in the positive z-direction. The Hamiltonian of the system can be split into two parts. The unperturbed Hamiltonian,

$$H_0 = \frac{p^2}{2\,m_e} + V(r), \tag{7.35}$$

and the perturbing Hamiltonian,

$$H_1 = e\,|\mathbf{E}|\,z. \tag{7.36}$$

Here, we are neglecting the small difference between the reduced mass, μ, and the electron mass, m_e.

It is assumed that the unperturbed energy eigenvalues and eigenstates are completely known. The electron spin is irrelevant in this problem (because the spin operators all commute with H_1), so we can ignore the spin degrees of freedom of the system. This implies that the system possesses no degenerate energy eigenvalues. Actually, this is not true for the $n \neq 1$ energy levels of the hydrogen atom, due to the special properties of a pure Coulomb potential. It is necessary to deal with this case separately, because the perturbation theory presented in Section 7.3 breaks down for degenerate unperturbed energy levels.

An energy eigenket of the unperturbed Hamiltonian is characterized by three quantum numbers—the radial quantum number n, and the two angular quantum numbers l and m (see Section 4.6). Let us denote such a ket $|n, l, m\rangle$, and let its energy be E_{nlm}. According to Equation (7.32), the change in this energy induced by a *small* electric field is given by

$$\varDelta E_{nlm} = e\,|\mathbf{E}|\,\langle n, l, m|\,z\,|n, l, m\rangle + e^2\,|\mathbf{E}|^2 \sum_{n',l',m'\neq n,l,m} \frac{|\langle n, l, m|\,z\,|n,'\,l', m'\rangle|^2}{E_{nlm} - E_{n'l'm'}}. \tag{7.37}$$

Now, since

$$L_z = x\,p_y - y\,p_x, \tag{7.38}$$

it follows that

$$[L_z, z] = 0. \tag{7.39}$$

Thus,

$$\langle n, l, m|\,[L_z, z]\,|n', l', m'\rangle = 0, \tag{7.40}$$

giving

$$(m - m')\,\langle n, l, m|\,z\,|n', l', m'\rangle = 0, \tag{7.41}$$

because $|n, l, m\rangle$ is, by definition, an eigenstate of L_z with eigenvalue $m\,\hbar$. It is clear, from the above relation, that the matrix element $\langle n, l, m|\,z\,|n', l', m'\rangle$ is zero unless $m' = m$. This is termed the *selection rule* for the quantum number m.

Let us now determine the selection rule for l. We have

$$\begin{aligned}
[L^2, z] &= [L_x^2, z] + [L_y^2, z] = L_x\,[L_x, z] + [L_x, z]\,L_x + L_y\,[L_y, z] + [L_y, z]\,L_y \\
&= \mathrm{i}\,\hbar\left(-L_x\,y - y\,L_x + L_y\,x + x\,L_y\right) = 2\,\mathrm{i}\,\hbar\,(L_y\,x - L_x\,y + \mathrm{i}\,\hbar\,z) \\
&= 2\,\mathrm{i}\,\hbar\,(L_y\,x - y\,L_x) = 2\,\mathrm{i}\,\hbar\,(x\,L_y - L_x\,y),
\end{aligned} \tag{7.42}$$

where use has been made of Equations (4.1)–(4.6). Similarly,

$$[L^2, y] = 2\,\mathrm{i}\,\hbar\,(L_x\,z - x\,L_z), \tag{7.43}$$

$$[L^2, x] = 2\,\mathrm{i}\,\hbar\,(y\,L_z - L_y\,z). \tag{7.44}$$

Thus,

$$[L^2,[L^2,z]] = 2\,\mathrm{i}\,\hbar\left(L^2, L_y\,x - L_x\,y + \mathrm{i}\,\hbar\,z\right) = 2\,\mathrm{i}\,\hbar\left(L_y\,[L^2,x] - L_x\,[L^2,y] + \mathrm{i}\,\hbar\,[L^2,z]\right)$$
$$= -4\,\hbar^2\,L_y\,(y\,L_z - L_y\,z) + 4\,\hbar^2\,L_x\,(L_x\,z - x\,L_z) - 2\,\hbar^2\,(L^2\,z - z\,L^2). \tag{7.45}$$

This reduces to

$$[L^2,[L^2,z]] = -\hbar^2\left[4\,(L_x\,x + L_y\,y + L_z\,z)\,L_z - 4\,(L_x^2 + L_y^2 + L_z^2)\,z + 2\,(L^2\,z - z\,L^2)\right]. \tag{7.46}$$

However, it is clear from Equations (4.1)–(4.3) that

$$L_x\,x + L_y\,y + L_z\,z = 0. \tag{7.47}$$

Hence, we obtain

$$[L^2,[L^2,z]] = 2\,\hbar^2\,(L^2\,z + z\,L^2), \tag{7.48}$$

which can be expanded to give

$$L^4\,z - 2\,L^2\,z\,L^2 + z\,L^4 - 2\,\hbar^2\,(L^2\,z + z\,L^2) = 0. \tag{7.49}$$

Equation (7.49) implies that

$$\langle n,l,m|\,L^4\,z - 2\,L^2\,z\,L^2 + z\,L^4 - 2\,\hbar^2\,(L^2\,z + z\,L^2)\,|n',l',m'\rangle = 0. \tag{7.50}$$

This expression yields

$$\left[l^2\,(l+1)^2 - 2\,l\,(l+1)\,l'\,(l'+1) + l'^{\,2}\,(l'+1)^2 - 2\,l\,(l+1) - 2\,l'\,(l'+1)\right]\langle n,l,m|\,z\,|n',l',m'\rangle = 0, \tag{7.51}$$

which reduces to

$$(l+l'+2)\,(l+l')\,(l-l'+1)\,(l-l'-1)\,\langle n,l,m|\,z\,|n',l',m'\rangle = 0. \tag{7.52}$$

According to the above formula, the matrix element $\langle n,l,m|\,z\,|n',l',m'\rangle$ vanishes unless $l = l' = 0$ or $l' = l \pm 1$. This matrix element can be written

$$\langle n,l,m|\,z\,|n',l',m'\rangle = \int\!\!\int\!\!\int dV'\,\psi^*_{nlm}(r',\theta',\varphi')\,r'\cos\theta'\,\psi_{n'm'l'}(r',\theta',\varphi'), \tag{7.53}$$

where $\psi_{nlm}(\mathbf{x}') = \langle\mathbf{x}'|n,l,m\rangle$. Recall, however, that the wavefunction of an $l = 0$ state is spherically symmetric (see Section 4.3): i.e., $\psi_{n00}(\mathbf{x}') = \psi_{n00}(r')$. It follows from Equation (7.53) that the matrix element vanishes by symmetry when $l = l' = 0$. In conclusion, the matrix element $\langle n,l,m|\,z\,|n',l',m'\rangle$ is zero unless $l' = l \pm 1$. This is the selection rule for the quantum number l.

Application of the selection rules to Equation (7.37) yields

$$\varDelta E_{nlm} = e^2\,|\mathbf{E}|^2\sum_{n'}\sum_{l'=l\pm1}\frac{|\langle n,l,m|\,z\,|n',l',m\rangle|^2}{E_{nlm} - E_{n'l'm}}. \tag{7.54}$$

Note that all of the terms in Equation (7.37) that vary linearly with the electric field-strength vanish by symmetry, according to the selection rules. Only those terms that vary *quadratically* with the field-strength survive. The electrical polarizability, α, of an atom is defined in terms of the electric-field induced energy-shift of a given atomic state as follows:

$$\varDelta E = -\frac{1}{2}\,\alpha\,|\mathbf{E}|^2. \tag{7.55}$$

Consider the ground state of a hydrogen atom. (Recall, that we cannot address the $n > 1$ excited states because they are degenerate, and our theory cannot handle this at present). The polarizability of this state is given by

$$\alpha = 2\,e^2 \sum_{n>1} \frac{|\langle 1,0,0|\,z\,|n,1,0\rangle|^2}{E_{n00} - E_{100}}. \tag{7.56}$$

Here, we have made use of the fact that $E_{n10} = E_{n00}$ for a hydrogen atom.

The sum in the above expression can be evaluated approximately by noting that [see Equation (4.120)]

$$E_{n00} = -\frac{e^2}{8\pi\,\epsilon_0\,a_0\,n^2} \tag{7.57}$$

for a hydrogen atom, where

$$a_0 = \frac{4\pi\,\epsilon_0\,\hbar^2}{m_e\,e^2} \tag{7.58}$$

is the Bohr radius. We can write

$$E_{n00} - E_{100} \geq E_{200} - E_{100} = \frac{3}{4}\frac{e^2}{8\pi\,\epsilon_0\,a_0}. \tag{7.59}$$

Thus,

$$\alpha < \frac{16}{3}\,4\pi\,\epsilon_0\,a_0 \sum_{n>1} |\langle 1,0,0|\,z\,|n,1,0\rangle|^2. \tag{7.60}$$

However,

$$\sum_{n>1} |\langle 1,0,0|\,z\,|n,1,0\rangle|^2 = \sum_{n',l',m'} \langle 1,0,0|\,z\,|n',l',m'\rangle\langle n',m',l'|\,z\,|1,0,0\rangle = \langle 1,0,0|\,z^2\,|1,0,0\rangle, \tag{7.61}$$

where we have made use of the fact that the wavefunctions of a hydrogen atom form a complete set. It is easily demonstrated from the actual form of the ground-state wavefunction that

$$\langle 1,0,0|\,z^2\,|1,0,0\rangle = a_0^2. \tag{7.62}$$

Thus, we conclude that

$$\alpha < \frac{16}{3}\,4\pi\,\epsilon_0\,a_0^3 \simeq 5.3\,4\pi\,\epsilon_0\,a_0^3. \tag{7.63}$$

The exact result is

$$\alpha = \frac{9}{2}\,4\pi\,\epsilon_0\,a_0^3 = 4.5\,4\pi\,\epsilon_0\,a_0^3. \tag{7.64}$$

It is possible to obtain this result, without recourse to perturbation theory, by solving Schrödinger's equation in parabolic coordinates.

7.5 Degenerate Perturbation Theory

Let us now consider systems in which the eigenstates of the unperturbed Hamiltonian, H_0, possess *degenerate* energy levels. It is always possible to represent degenerate energy eigenstates as the simultaneous eigenstates of the Hamiltonian and some other Hermitian operator (or group of operators). Let us denote this operator (or group of operators) L. We can write

$$H_0\,|n,l\rangle = E_n\,|n,l\rangle, \tag{7.65}$$

and

$$L\,|n,l\rangle = L_{n\,l}\,|n,l\rangle, \tag{7.66}$$

where $[H_0, L] = 0$. Here, the E_n and the $L_{n\,l}$ are real numbers that depend on the quantum numbers n, and n and l, respectively. It is always possible to find a sufficient number of operators which commute with the Hamiltonian in order to ensure that the $L_{n\,l}$ are all different. In other words, we can choose L such that the quantum numbers n and l *uniquely* specify each eigenstate. Suppose that for each value of n there are N_n different values of l: i.e., the nth energy eigenstate is N_n-fold degenerate.

In general, L *does not* commute with the perturbing Hamiltonian, H_1. This implies that the modified energy eigenstates are *not* eigenstates of L. In this situation, we expect the perturbation to split the degeneracy of the energy levels, so that each modified eigenstate $|n,l\rangle'$ acquires a unique energy eigenvalue E'_{nl}. Let us naively attempt to use the standard perturbation theory of Section 7.3 to evaluate the modified energy eigenstates and energy levels. A direct generalization of Equations (7.32) and (7.33) yields

$$E'_{nl} = E_n + e_{nlnl} + \sum_{n',l'\neq n,l} \frac{|e_{n'l'nl}|^2}{E_n - E_{n'}} + O(\epsilon^3), \tag{7.67}$$

and

$$|n,l\rangle' = |n,l\rangle + \sum_{n',l'\neq n,l} \frac{e_{n'l'nl}}{E_n - E_{n'}}\,|n',l'\rangle + O(\epsilon^2), \tag{7.68}$$

where

$$e_{n'l'nl} = \langle n',l'|\,H_1\,|n,l\rangle. \tag{7.69}$$

It is fairly obvious that the summations in Equations (7.67) and (7.68) are not well-behaved if the nth energy level is degenerate. The problem terms are those involving unperturbed eigenstates labeled by the same value of n, but different values of l: i.e., those states whose unperturbed energies are E_n. These terms give rise to singular factors $1/(E_n - E_n)$ in the summations. Note, however, that this problem would not exist if the matrix elements, $e_{nl'nl}$, of the perturbing Hamiltonian between distinct, degenerate, unperturbed energy eigenstates corresponding to the eigenvalue E_n were zero. In other words, if

$$\langle n,l'|\,H_1\,|n,l\rangle = \lambda_{n\,l}\,\delta_{l\,l'}, \tag{7.70}$$

then all of the singular terms in Equations (7.67) and (7.68) would vanish.

In general, Equation (7.70) is not satisfied. Fortunately, we can always redefine the unperturbed energy eigenstates belonging to the eigenvalue E_n in such a manner that Equation (7.70) is satisfied.

Let us define N_n new states that are linear combinations of the N_n original degenerate eigenstates corresponding to the eigenvalue E_n:

$$|n, l^{(1)}\rangle = \sum_{k=1,N_n} \langle n, k|n, l^{(1)}\rangle |n, k\rangle. \tag{7.71}$$

Note that these new states are also degenerate energy eigenstates of the unperturbed Hamiltonian corresponding to the eigenvalue E_n. The $|n, l^{(1)}\rangle$ are chosen in such a manner that they are eigenstates of the perturbing Hamiltonian, H_1. Thus,

$$H_1\,|n, l^{(1)}\rangle = \lambda_{n\,l}\,|n, l^{(1)}\rangle. \tag{7.72}$$

The $|n, l^{(1)}\rangle$ are also chosen so that they are orthogonal, and have unit lengths. It follows that

$$\langle n, l'^{(1)}|\,H_1\,|n, l^{(1)}\rangle = \lambda_{n\,l}\,\delta_{l\,l'}. \tag{7.73}$$

Thus, if we use the new eigenstates, instead of the old ones, then we can employ Equations (7.67) and (7.68) directly, because all of the singular terms vanish. The only remaining difficulty is to determine the new eigenstates in terms of the original ones.

Now

$$\sum_{l=1,N_n} |n, l\rangle\langle n, l| = 1, \tag{7.74}$$

where 1 denotes the identity operator in the sub-space of all unperturbed energy eigenkets corresponding to the eigenvalue E_n. Using this completeness relation, the operator eigenvalue equation (7.72) can be transformed into a straightforward matrix eigenvalue equation:

$$\sum_{l''=1,N_n} \langle n, l'|H_1|n, l''\rangle\langle n, l''|n, l^{(1)}\rangle = \lambda_{n\,l}\,\langle n, l'|n, l^{(1)}\rangle. \tag{7.75}$$

This can be written more transparently as

$$\mathbf{U}\,\mathbf{x} = \lambda\,\mathbf{x}, \tag{7.76}$$

where the elements of the $N_n \times N_n$ Hermitian matrix $\mathbf{U}$ are

$$U_{j\,k} = \langle n, j|\,H_1\,|n, k\rangle. \tag{7.77}$$

Provided that the determinant of $\mathbf{U}$ is non-zero, Equation (7.76) can always be solved to give N_n eigenvalues $\lambda_{n\,l}$ (for $l = 1$ to N_n), with N_n corresponding eigenvectors $\mathbf{x}_{n\,l}$. The eigenvectors specify the weights of the new eigenstates in terms of the original eigenstates: i.e.,

$$(\mathbf{x}_{n\,l})_k = \langle n, k|n, l^{(1)}\rangle, \tag{7.78}$$

for $k = 1$ to N_n. In our new scheme, Equations (7.67) and (7.68) yield

$$E'_{nl} = E_n + \lambda_{n\,l} + \sum_{n'\neq n,l'} \frac{|e_{n'l'nl}|^2}{E_n - E_{n'}} + O(\epsilon^3), \tag{7.79}$$

and

$$|n, l^{(1)}\rangle' = |n, l^{(1)}\rangle + \sum_{n'\neq n, l'} \frac{e_{n'l'nl}}{E_n - E_{n'}}\,|n', l'\rangle + O(\epsilon^2). \tag{7.80}$$

There are no singular terms in these expressions, because the summations are over $n' \neq n$: i.e., they specifically exclude the problematic, degenerate, unperturbed energy eigenstates corresponding to the eigenvalue E_n. Note that the first-order energy-shifts are equivalent to the eigenvalues of the matrix equation, (7.76).

7.6 Linear Stark Effect

Let us examine the effect of an electric field on the excited energy levels of a hydrogen atom. For instance, consider the $n = 2$ states. There is a single $l = 0$ state, usually referred to as $2s$, and three $l = 1$ states (with $m = -1, 0, 1$), usually referred to as $2p$. All of these states possess the same energy, $E_{200} = -e^2/(32\pi\,\epsilon_0\,a_0)$. As in Section 7.4, the perturbing Hamiltonian is

$$H_1 = e\,|\mathbf{E}|\,z. \tag{7.81}$$

In order to apply perturbation theory, we have to solve the matrix eigenvalue equation

$$\mathbf{U}\,\mathbf{x} = \lambda\,\mathbf{x}, \tag{7.82}$$

where $\mathbf{U}$ is the array of the matrix elements of H_1 between the degenerate $2s$ and $2p$ states. Thus,

$$\mathbf{U} = e\,|\mathbf{E}| \begin{pmatrix} 0 & \langle 2,0,0|\,z\,|2,1,0\rangle & 0 & 0 \\ \langle 2,1,0|\,z\,|2,0,0\rangle & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{7.83}$$

where the rows and columns correspond to the $|2,0,0\rangle$, $|2,1,0\rangle$, $|2,1,1\rangle$, and $|2,1,-1\rangle$ states, respectively. Here, we have made use of the selection rules, which tell us that the matrix element of z between two hydrogen atom states is zero unless the states possess the same m quantum number, and l quantum numbers that differ by unity. It is easily demonstrated, from the exact forms of the $2s$ and $2p$ wavefunctions, that

$$\langle 2,0,0|\,z\,|2,1,0\rangle = \langle 2,1,0|\,z\,|2,0,0\rangle = 3\,a_0. \tag{7.84}$$

It can be seen, by inspection, that the eigenvalues of $\mathbf{U}$ are $\lambda_1 = 3\,e\,a_0\,|\mathbf{E}|$, $\lambda_2 = -3\,e\,a_0\,|\mathbf{E}|$,

$\lambda_3 = 0$, and $\lambda_4 = 0$. The corresponding eigenvectors are

$$\mathbf{x}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \\ 0 \end{pmatrix}, \tag{7.85}$$

$$\mathbf{x}_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \\ 0 \end{pmatrix}, \tag{7.86}$$

$$\mathbf{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \tag{7.87}$$

$$\mathbf{x}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{7.88}$$

It follows from Section 7.5 that the simultaneous eigenstates of the unperturbed Hamiltonian and the perturbing Hamiltonian take the form

$$|1\rangle = \frac{|2,0,0\rangle + |2,1,0\rangle}{\sqrt{2}}, \tag{7.89}$$

$$|2\rangle = \frac{|2,0,0\rangle - |2,1,0\rangle}{\sqrt{2}}, \tag{7.90}$$

$$|3\rangle = |2,1,1\rangle, \tag{7.91}$$

$$|4\rangle = |2,1,-1\rangle. \tag{7.92}$$

In the absence of an electric field, all of these states possess the same energy, E_{200}. The first-order energy-shifts induced by an electric field are given by

$$\Delta E_1 = +3\, e\, a_0\, |\mathbf{E}|, \tag{7.93}$$

$$\Delta E_2 = -3\, e\, a_0\, |\mathbf{E}|, \tag{7.94}$$

$$\Delta E_3 = 0, \tag{7.95}$$

$$\Delta E_4 = 0. \tag{7.96}$$

Thus, the energies of states 1 and 2 are shifted upwards and downwards, respectively, by an amount $3\, e\, a_0\, |\mathbf{E}|$ in the presence of an electric field. States 1 and 2 are orthogonal linear combinations of the original $2s$ and $2p(m = 0)$ states. Note that the energy-shifts are *linear* in the electric field-strength, so this is a much larger effect that the quadratic effect described in Section 7.4. The energies of states 3 and 4 (which are equivalent to the original $2p(m = 1)$ and $2p(m = -1)$ states, respectively) are not affected to first order. Of course, to second order the energies of these states are shifted by an amount that depends on the square of the electric field-strength.

Note that the linear Stark effect depends crucially on the degeneracy of the $2s$ and $2p$ states. This degeneracy is a special property of a pure Coulomb potential, and, therefore, only applies to a hydrogen atom. Thus, alkali metal atoms do not exhibit the linear Stark effect.

7.7 Fine Structure

Let us now consider the energy levels of hydrogen-like atoms (i.e., alkali metal atoms) in more detail. The outermost electron moves in a spherically symmetric potential $V(r)$ due to the nuclear charge and the charges of the other electrons (which occupy spherically symmetric closed shells). The shielding effect of the inner electrons causes $V(r)$ to depart from the pure Coulomb form. This splits the degeneracy of states characterized by the same value of n, but different values of l. In fact, higher l states have higher energies.

Let us examine a phenomenon known as *fine structure*, which is due to interaction between the spin and orbital angular momenta of the outermost electron. This electron experiences an electric field

$$\mathbf{E} = \frac{\nabla V}{e}. \tag{7.97}$$

However, a non-relativistic charge moving in an electric field also experiences an effective magnetic field

$$\mathbf{B} = -\frac{\mathbf{v}\times\mathbf{E}}{c^2}. \tag{7.98}$$

Now, an electron possesses a spin magnetic moment [see Equation (5.46)]

$$\boldsymbol{\mu} = -\frac{e\,\mathbf{S}}{m_e}. \tag{7.99}$$

We, therefore, expect a spin-orbit contribution to the Hamiltonian of the form

$$H_{LS} = -\boldsymbol{\mu}\cdot\mathbf{B} = -\frac{e\,\mathbf{S}}{m_e\,c^2}\cdot\mathbf{v}\times\left(\frac{1}{e}\frac{\mathbf{x}}{r}\frac{dV}{dr}\right) = \frac{1}{m_e^2\,c^2\,r}\frac{dV}{dr}\,\mathbf{L}\cdot\mathbf{S}, \tag{7.100}$$

where $\mathbf{L} = m_e\,\mathbf{x}\times\mathbf{v}$ is the orbital angular momentum. Actually, when the above expression is compared to the observed spin-orbit interaction, it is found to be too large by a factor of two. There is a classical explanation for this, due to spin precession, which we need not go into. The correct quantum mechanical explanation requires a relativistically covariant treatment of electron dynamics (this is achieved using the so-called *Dirac equation*—see Chapter 11).

Let us now apply perturbation theory to a hydrogen-like atom, using H_{LS} as the perturbation (with H_{LS} taking one half of the value given above), and

$$H_0 = \frac{p^2}{2\,m_e} + V(r) \tag{7.101}$$

as the unperturbed Hamiltonian. We have two choices for the energy eigenstates of H_0. We can adopt the simultaneous eigenstates of H_0, L^2, S^2, L_z and S_z, or the simultaneous eigenstates of H_0, L^2, S^2, J^2, and J_z, where $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is the total angular momentum. Although the departure of $V(r)$ from a pure $1/r$ form splits the degeneracy of same n, different l, states, those states characterized by the same values of n and l, but different values of m_l, are still degenerate. (Here, m_l, m_s, and m_j are the quantum numbers corresponding to L_z, S_z, and J_z, respectively.) Moreover, with the addition of spin degrees of freedom, each state is doubly degenerate due to the two possible orientations of the electron spin (i.e., $m_s = \pm 1/2$). Thus, we are still dealing with a highly degenerate system. We know, from Section 7.6, that the application of perturbation theory to a degenerate system is greatly simplified if the basis eigenstates of the unperturbed Hamiltonian are also eigenstates of the perturbing Hamiltonian. Now, the perturbing Hamiltonian, H_{LS}, is proportional to $\mathbf{L}\cdot\mathbf{S}$, where

$$\mathbf{L}\cdot\mathbf{S} = \frac{J^2 - L^2 - S^2}{2}. \tag{7.102}$$

It is fairly obvious that the first group of operators (H_0, L^2, S^2, L_z and S_z) *does not* commute with H_{LS}, whereas the second group (H_0, L^2, S^2, J^2, and J_z) does. In fact, $\mathbf{L}\cdot\mathbf{S}$ is just a combination of operators appearing in the second group. Thus, it is advantageous to work in terms of the eigenstates of the second group of operators, rather than those of the first group.

We now need to find the simultaneous eigenstates of H_0, L^2, S^2, J^2, and J_z. This is equivalent to finding the eigenstates of the total angular momentum resulting from the addition of two angular momenta: $j_1 = l$, and $j_2 = s = 1/2$. According to Equation (6.26), the allowed values of the total angular momentum are $j = l + 1/2$ and $j = l - 1/2$. We can write

$$|l + 1/2, m\rangle = \cos\alpha\,|m - 1/2, 1/2\rangle + \sin\alpha\,|m + 1/2, -1/2\rangle, \tag{7.103}$$

$$|l - 1/2, m\rangle = -\sin\alpha\,|m - 1/2, 1/2\rangle + \cos\alpha\,|m + 1/2, -1/2\rangle. \tag{7.104}$$

Here, the kets on the left-hand side are $|j, m_j\rangle$ kets, whereas those on the right-hand side are $|m_l, m_s\rangle$ kets (the j_1, j_2 labels have been dropped, for the sake of clarity). We have made use of the fact that the Clebsch-Gordon coefficients are automatically zero unless $m_j = m_l + m_s$. We have also made use of the fact that both the $|j, m_j\rangle$ and $|m_l, m_s\rangle$ kets are orthonormal, and have unit lengths. We now need to determine

$$\cos\alpha = \langle m - 1/2, 1/2|l + 1/2, m\rangle, \tag{7.105}$$

where the Clebsch-Gordon coefficient is written in $\langle m_l, m_s|j, m_j\rangle$ form.

Let us now employ the recursion relation for Clebsch-Gordon coefficients, Equation (6.32), with $j_1 = l, j_2 = 1/2, j = l + 1/2, m_1 = m - 1/2, m_2 = 1/2$ (lower sign). We obtain

$$\begin{aligned}&[(l + 1/2)\,(l + 3/2) - m\,(m + 1)]^{1/2}\,\langle m - 1/2, 1/2|l + 1/2, m\rangle\\ &\quad = [l\,(l + 1) - (m - 1/2)\,(m + 1/2)]^{1/2}\,\langle m + 1/2, 1/2|l + 1/2, m + 1\rangle,\end{aligned} \tag{7.106}$$

which reduces to

$$\langle m - 1/2, 1/2|l + 1/2, m\rangle = \sqrt{\frac{l + m + 1/2}{l + m + 3/2}}\,\langle m + 1/2, 1/2|l + 1/2, m + 1\rangle. \tag{7.107}$$

We can use this formula to successively increase the value of m_l. For instance,

$$\langle m-1/2, 1/2|l+1/2, m\rangle = \sqrt{\frac{l+m+1/2}{l+m+3/2}}\sqrt{\frac{l+m+3/2}{l+m+5/2}}\langle m+3/2, 1/2|l+1/2, m+2\rangle. \quad (7.108)$$

This procedure can be continued until m_l attains its maximum possible value, l. Thus,

$$\langle m-1/2, 1/2|l+1/2, m\rangle = \sqrt{\frac{l+m+1/2}{2\,l+1}}\,\langle l, 1/2|l+1/2, l+1/2\rangle. \quad (7.109)$$

Consider the situation in which m_l and m both take their maximum values, l and $1/2$, respectively. The corresponding value of m_j is $l+1/2$. This value is possible when $j = l+1/2$, but not when $j = l-1/2$. Thus, the $|m_l, m_s\rangle$ ket $|l, 1/2\rangle$ must be equal to the $|j, m_j\rangle$ ket $|l+1/2, l+1/2\rangle$, up to an arbitrary phase-factor. By convention, this factor is taken to be unity, giving

$$\langle l, 1/2|l+1/2, l+1/2\rangle = 1. \quad (7.110)$$

It follows from Equation (7.109) that

$$\cos\alpha = \langle m-1/2, 1/2|l+1/2, m\rangle = \sqrt{\frac{l+m+1/2}{2\,l+1}}. \quad (7.111)$$

Hence,

$$\sin^2\alpha = 1 - \frac{l+m+1/2}{2\,l+1} = \frac{l-m+1/2}{2\,l+1}. \quad (7.112)$$

We now need to determine the sign of $\sin\alpha$. A careful examination of the recursion relation, Equation (6.32), shows that the plus sign is appropriate. Thus,

$$|l+1/2, m\rangle = \sqrt{\frac{l+m+1/2}{2\,l+1}}\,|m-1/2, 1/2\rangle + \sqrt{\frac{l-m+1/2}{2\,l+1}}\,|m+1/2, -1/2\rangle, \quad (7.113)$$

$$|l-1/2, m\rangle = -\sqrt{\frac{l-m+1/2}{2\,l+1}}\,|m-1/2, 1/2\rangle + \sqrt{\frac{l+m+1/2}{2\,l+1}}\,|m+1/2, -1/2\rangle. \quad (7.114)$$

It is convenient to define so called *spin-angular functions* using the Pauli two-component formalism:

$$\begin{aligned}\mathcal{Y}_{l\,m}^{j=l\pm 1/2} \equiv \mathcal{Y}_{l\,m}^{\pm} &= \pm\sqrt{\frac{l\pm m+1/2}{2\,l+1}}\;Y_{l\;m-1/2}(\theta,\varphi)\,\chi_+ + \sqrt{\frac{l\mp m+1/2}{2\,l+1}}\;Y_{l\;m+1/2}(\theta,\varphi)\,\chi_- \\ &= \frac{1}{\sqrt{2\,l+1}}\begin{pmatrix} \pm\sqrt{l\pm m+1/2}\;Y_{l\;m-1/2}(\theta,\varphi) \\ \sqrt{l\mp m+1/2}\;Y_{l\;m+1/2}(\theta,\varphi)\end{pmatrix}. \end{aligned} \quad (7.115)$$

These functions are eigenfunctions of the total angular momentum for spin one-half particles, just as the spherical harmonics are eigenfunctions of the orbital angular momentum. A general wavefunction for an energy eigenstate in a hydrogen-like atom is written

$$\psi_{nlm\pm} = R_{n\,l}(r)\,\mathcal{Y}_{l\,m}^{\pm}. \quad (7.116)$$

The radial part of the wavefunction, $R_{nl}(r)$, depends on the radial quantum number n and the angular quantum number l. The wavefunction is also labeled by m, which is the quantum number associated with J_z. For a given choice of l, the quantum number j (i.e., the quantum number associated with J^2) can take the values $l \pm 1/2$.

The $|l \pm 1/2, m\rangle$ kets are eigenstates of $\mathbf{L}\cdot\mathbf{S}$, according to Equation (7.102). Thus,

$$\mathbf{L}\cdot\mathbf{S}\,|j = l \pm 1/2, m_j = m\rangle = \frac{\hbar^2}{2}\,[\,j\,(j+1) - l\,(l+1) - 3/4]\,|j, m\rangle, \tag{7.117}$$

giving

$$\mathbf{L}\cdot\mathbf{S}\,|l + 1/2, m\rangle = \frac{l\,\hbar^2}{2}\,|l + 1/2, m\rangle, \tag{7.118}$$

$$\mathbf{L}\cdot\mathbf{S}\,|l - 1/2, m\rangle = -\frac{(l+1)\,\hbar^2}{2}\,|l - 1/2, m\rangle. \tag{7.119}$$

It follows that

$$\oint d\Omega\,(\mathcal{Y}^{+}_{l\,m})^{\dagger}\,\mathbf{L}\cdot\mathbf{S}\,\mathcal{Y}^{+}_{l\,m} = \frac{l\,\hbar^2}{2}, \tag{7.120}$$

$$\oint d\Omega\,(\mathcal{Y}^{-}_{l\,m})^{\dagger}\,\mathbf{L}\cdot\mathbf{S}\,\mathcal{Y}^{-}_{l\,m} = -\frac{(l+1)\,\hbar^2}{2}, \tag{7.121}$$

where the integrals are over all solid angle, $d\Omega = \sin\theta\,d\theta\,d\varphi$.

Let us now apply degenerate perturbation theory to evaluate the shift in energy of a state whose wavefunction is $\psi_{nlm\pm}$ due to the spin-orbit Hamiltonian, H_{LS}. To first order, the energy-shift is given by

$$\Delta E_{nlm\pm} = \int dV\,(\psi_{nlm\pm})^{\dagger}\,H_{LS}\,\psi_{nlm\pm}, \tag{7.122}$$

where the integral is over all space, $dV = r^2\,d\Omega$. Equations (7.100) (remembering the factor of two), (7.116), and (7.120)–(7.121) yield

$$\Delta E_{nlm+} = +\frac{1}{2\,m_e^2\,c^2}\left\langle\frac{1}{r}\frac{dV}{dr}\right\rangle\frac{l\,\hbar^2}{2}, \tag{7.123}$$

$$\Delta E_{nlm-} = -\frac{1}{2\,m_e^2\,c^2}\left\langle\frac{1}{r}\frac{dV}{dr}\right\rangle\frac{(l+1)\,\hbar^2}{2}, \tag{7.124}$$

where

$$\left\langle\frac{1}{r}\frac{dV}{dr}\right\rangle = \int_0^\infty dr\,r^2\,(R_{n\,l})^*\,\frac{1}{r}\frac{dV}{dr}\,R_{n\,l}. \tag{7.125}$$

Let us now apply the above result to the case of a sodium atom. In chemist's notation, the ground state is written

$$(1s)^2(2s)^2(2p)^6(3s). \tag{7.126}$$

The inner ten electrons effectively form a spherically symmetric electron cloud. We are interested in the excitation of the eleventh electron from $3s$ to some higher energy state. The closest (in energy) unoccupied state is $3p$. This state has a higher energy than $3s$ due to the deviations of the potential from the pure Coulomb form. In the absence of spin-orbit interaction, there are six degenerate $3p$ states. The spin-orbit interaction breaks the degeneracy of these states. The modified states are labeled $(3p)_{1/2}$ and $(3p)_{3/2}$, where the subscript refers to the value of j. The four $(3p)_{3/2}$ states lie at a slightly higher energy level than the two $(3p)_{1/2}$ states, because the radial integral (7.125) is positive. The splitting of the $(3p)$ energy levels of the sodium atom can be observed using a spectroscope. The well-known sodium D line is associated with transitions between the $3p$ and $3s$ states. The fact that there are two slightly different $3p$ energy levels (note that spin-orbit coupling does not split the $3s$ energy levels) means that the sodium D line actually consists of two very closely spaced spectroscopic lines. It is easily demonstrated that the ratio of the typical spacing of Balmer lines to the splitting brought about by spin-orbit interaction is about $1 : \alpha^2$, where

$$\alpha = \frac{e^2}{2\,\epsilon_0\,h\,c} = \frac{1}{137} \tag{7.127}$$

is the fine structure constant. Actually, Equations (7.123)–(7.124) are not entirely correct, because we have neglected an effect (namely, the relativistic mass correction of the electron) that is the same order of magnitude as spin-orbit coupling. (See Exercises 7.3 and 7.4.)

7.8 Zeeman Effect

Consider a hydrogen-like atom placed in a uniform z-directed magnetic field. The change in energy of the outermost electron is

$$H_B = -\boldsymbol{\mu}\cdot\mathbf{B}, \tag{7.128}$$

where

$$\boldsymbol{\mu} = -\frac{e}{2\,m_e}\,(\mathbf{L} + 2\,\mathbf{S}) \tag{7.129}$$

is its magnetic moment, including both the spin and orbital contributions. Thus,

$$H_B = \frac{e\,B}{2\,m_e}\,(L_z + 2\,S_z). \tag{7.130}$$

Suppose that the energy-shifts induced by the magnetic field are much smaller than those induced by spin-orbit interaction. In this situation, we can treat H_B as a small perturbation acting on the eigenstates of $H_0 + H_{LS}$. Of course, these states are the simultaneous eigenstates of J^2 and J_z. Let us consider one of these states, labeled by the quantum numbers j and m, where $j = l \pm 1/2$. From standard perturbation theory, the first-order energy-shift in the presence of a magnetic field is

$$\Delta E_{nlm\pm} = \langle l \pm 1/2, m|\,H_B\,|l \pm 1/2, m\rangle. \tag{7.131}$$

Because

$$L_z + 2\,S_z = J_z + S_z, \tag{7.132}$$

we find that

$$\Delta E_{nlm\pm} = \frac{e\,B}{2\,m_e}\,(m\,\hbar + \langle l \pm 1/2, m|\,S_z\,|l \pm 1/2, m\rangle\,)\,. \tag{7.133}$$

Now, from Equations (7.113)–(7.114),

$$|l \pm 1/2, m\rangle = \pm\sqrt{\frac{l \pm m + 1/2}{2\,l+1}}\,|m - 1/2, 1/2\rangle + \sqrt{\frac{l \mp m + 1/2}{2\,l+1}}\,|m + 1/2, -1/2\rangle. \tag{7.134}$$

It follows that

$$\langle l \pm 1/2, m|\,S_z\,|l \pm 1/2, m\rangle = \frac{\hbar}{2\,(2\,l+1)}\,[(l \pm m + 1/2) - (l \mp m + 1/2)] = \pm\frac{m\,\hbar}{2\,l+1}. \tag{7.135}$$

Thus, we obtain the so-called *Lande formula* for the energy-shift induced by a weak magnetic field:

$$\Delta E_{nlm\pm} = \frac{e\,\hbar\,B}{2\,m_e}\,m\left(1 \pm \frac{1}{2\,l+1}\right). \tag{7.136}$$

Let us apply this theory to the sodium atom. We have already seen that the non-Coulomb potential splits the degeneracy of the $3s$ and $3p$ states, the latter states acquiring a higher energy. The spin-orbit interaction splits the six $3p$ states into two groups, with four $j = 3/2$ states lying at a slightly higher energy than two $j = 1/2$ states. According to Equation (7.136), a magnetic field splits the $(3p)_{3/2}$ quadruplet of states, each state acquiring a different energy. In fact, the energy of each state becomes dependent on the quantum number m, which measures the projection of the total angular momentum along the z-axis. States with higher m values have higher energies. A magnetic field also splits the $(3p)_{1/2}$ doublet of states. However, it is evident from Equation (7.136) that these states are split by a lesser amount than the $j = 3/2$ states.

Suppose that we increase the strength of the magnetic field, so that the energy-shift due to the magnetic field becomes comparable to the energy-shift induced by spin-orbit interaction. Clearly, in this situation, it does not make much sense to think of H_B as a small interaction term operating on the eigenstates of $H_0 + H_{LS}$. In fact, this intermediate case is very difficult to analyze. Let us, instead, consider the extreme limit in which the energy-shift due to the magnetic field greatly exceeds that induced by spin-orbit effects. This is called the *Paschen-Back limit*.

In the Paschen-Back limit, we can think of the spin-orbit Hamiltonian, H_{LS}, as a small interaction term operating on the eigenstates of $H_0 + H_B$. Note that the magnetic Hamiltonian, H_B, commutes with L^2, S^2, L_z, S_z, but does not commute with L^2, S^2, J^2, J_z. Thus, in an intense magnetic field, the energy eigenstates of a hydrogen-like atom are approximate eigenstates of the spin and orbital angular momenta, but are not eigenstates of the total angular momentum. We can label each state by the quantum numbers n (the energy quantum number), l, m_l, and m_s. Thus, our energy eigenkets are written $|n, l, m_l, m_s\rangle$. The unperturbed Hamiltonian, H_0, causes states with different values of the quantum numbers n and l to have different energies. However, states with the same value of n and l, but different values of m_l and m_s, are degenerate. The shift in energy due to the magnetic field is simply

$$\Delta E_{n\,l\,m_l\,m_s} = \langle n, l, m_l, m_s|\,H_B\,|n, l, m_l, m_s\rangle = \frac{e\,\hbar\,B}{2\,m_e}\,(m_l + 2\,m_s). \tag{7.137}$$

Thus, states with different values of $m_l + 2\,m_s$ acquire different energies.

Let us apply this result to a sodium atom. In the absence of a magnetic field, the six $3p$ states form two groups of four and two states, depending on the values of their total angular momentum. In the presence of an intense magnetic field, the $3p$ states are split into five groups. There is a state with $m_l + 2\,m_s = 2$, a state with $m_l + 2\,m_s = 1$, two states with $m_l + 2\,m_s = 0$, a state with $m_l + 2\,m_s = -1$, and a state with $m_l + 2\,m_s = -2$. These groups are equally spaced in energy, the energy difference between adjacent groups being $e\,\hbar\,B/2\,m_e$.

The energy-shift induced by the spin-orbit Hamiltonian is given by

$$\Delta E_{n\,l\,m_l\,m_s} = \langle n, l, m_l, m_s|\,H_{LS}\,|n, l, m_l, m_s\rangle, \tag{7.138}$$

where

$$H_{LS} = \frac{1}{2\,m_e^2\,c^2}\,\frac{1}{r}\,\frac{dV}{dr}\,\mathbf{L}\cdot\mathbf{S}. \tag{7.139}$$

Now,

$$\langle\mathbf{L}\cdot\mathbf{S}\rangle = \langle\,L_z\,S_z + (L^+\,S^- + L^-\,S^+)/2\,\rangle = \hbar^2\,m_l\,m_s, \tag{7.140}$$

since

$$\langle L^\pm\rangle = \langle S^\pm\rangle = 0 \tag{7.141}$$

for expectation values taken between the simultaneous eigenkets of L_z and S_z. Thus,

$$\Delta E_{n\,l\,m_l\,m_s} = \frac{\hbar^2\,m_l\,m_s}{2\,m_e^2\,c^2}\left\langle\frac{1}{r}\frac{dV}{dr}\right\rangle. \tag{7.142}$$

Let us apply the above result to a sodium atom. In the presence of an intense magnetic field, the $3p$ states are split into five groups with (m_l, m_s) quantum numbers $(1, 1/2)$, $(0, 1/2)$, $(1, -1/2)$ or $(-1, 1/2)$, $(0, -1/2)$, and $(-1, -1/2)$, respectively, in order of decreasing energy. The spin-orbit term increases the energy of the highest energy state, does not affect the next highest energy state, decreases, but does not split, the energy of the doublet, does not affect the next lowest energy state, and increases the energy of the lowest energy state. The net result is that the five groups of states are no longer equally spaced in energy.

The typical magnetic field-strength needed to access the Paschen-Bach limit is

$$B_{PB} \sim \alpha^2\,\frac{e\,m_e}{\epsilon_0\,h\,a_0} \simeq 25\ \text{ tesla}. \tag{7.143}$$

7.9 Hyperfine Structure

The proton in a hydrogen atom is a spin one-half charged particle, and therefore possesses a magnetic moment. By analogy with Equation (7.99), we can write

$$\boldsymbol{\mu}_p = \frac{g_p\,e}{2\,m_p}\,\mathbf{S}_p, \tag{7.144}$$

where $\boldsymbol{\mu}_p$ is the proton magnetic moment, $\mathbf{S}_p$ is the proton spin, and the proton gyromagnetic ratio g_p is found experimentally to take that value 5.59. Note that the magnetic moment of a proton is much smaller (by a factor of order m_e/m_p) than that of an electron. According to classical electromagnetism, the proton's magnetic moment generates a magnetic field of the form

$$\mathbf{B} = \frac{\mu_0}{4\pi\, r^3}\left[3\,(\boldsymbol{\mu}_p\cdot\mathbf{e}_r)\,\mathbf{e}_r - \boldsymbol{\mu}_p\right] + \frac{2\,\mu_0}{3}\,\boldsymbol{\mu}_p\,\delta(\mathbf{x}), \tag{7.145}$$

where $\mathbf{e}_r = \mathbf{x}/r$, and $r = |\mathbf{x}|$. We can understand the origin of the delta-function term in the above expression by thinking of the proton as a tiny current loop centred on the origin. All magnetic field-lines generated by the loop must pass through the loop. Hence, if the size of the loop goes to zero then the field will be infinite at the origin, and this contribution is what is reflected by the delta-function term. Now, the Hamiltonian of the electron in the magnetic field generated by the proton is simply

$$H_1 = -\boldsymbol{\mu}_e\cdot\mathbf{B}, \tag{7.146}$$

where

$$\boldsymbol{\mu}_e = -\frac{e}{m_e}\,\mathbf{S}_e. \tag{7.147}$$

Here, $\boldsymbol{\mu}_e$ is the electron magnetic moment [see Equation (7.99)], and $\mathbf{S}_e$ the electron spin. Thus, the perturbing Hamiltonian is written

$$H_1 = \frac{\mu_0\,g_p\,e^2}{8\pi\,m_p\,m_e}\left(\frac{3\,(\mathbf{S}_p\cdot\mathbf{e}_r)\,(\mathbf{S}_e\cdot\mathbf{e}_r) - \mathbf{S}_p\cdot\mathbf{S}_e}{r^3}\right) + \frac{\mu_0\,g_p\,e^2}{3\,m_p\,m_e}\,\mathbf{S}_p\cdot\mathbf{S}_e\,\delta(\mathbf{x}). \tag{7.148}$$

Note that, because we have neglected coupling between the proton spin and the magnetic field generated by the electron's orbital motion, the above expression is only valid for $l = 0$ states.

According to standard first-order perturbation theory, the energy-shift induced by spin-spin coupling between the proton and the electron is the expectation value of the perturbing Hamiltonian. Hence,

$$\varDelta E = \frac{\mu_0\,g_p\,e^2}{8\pi\,m_p\,m_e}\left\langle\frac{3\,(\mathbf{S}_p\cdot\mathbf{e}_r)\,(\mathbf{S}_e\cdot\mathbf{e}_r) - \mathbf{S}_p\cdot\mathbf{S}_e}{r^3}\right\rangle + \frac{\mu_0\,g_p\,e^2}{3\,m_p\,m_e}\,\langle\mathbf{S}_p\cdot\mathbf{S}_e\rangle\,|\psi(0)|^2. \tag{7.149}$$

For the ground-state of hydrogen, which is spherically symmetric, the first term in the above expression vanishes by symmetry. Moreover, it is easily demonstrated that $|\psi_{000}(0)|^2 = 1/(\pi\,a_0^3)$. Thus, we obtain

$$\varDelta E_{000} = \frac{\mu_0\,g_p\,e^2}{3\pi\,m_p\,m_e\,a_0^3}\,\langle\mathbf{S}_p\cdot\mathbf{S}_e\rangle. \tag{7.150}$$

Let

$$\mathbf{S} = \mathbf{S}_e + \mathbf{S}_p \tag{7.151}$$

be the total spin. We can show that

$$\mathbf{S}_p\cdot\mathbf{S}_e = \frac{1}{2}\,(S^2 - S_e^{\,2} - S_p^{\,2}). \tag{7.152}$$

Thus, the simultaneous eigenstates of the perturbing Hamiltonian and the main Hamiltonian are the simultaneous eigenstates of $S_e^{\,2}$, $S_p^{\,2}$, and S^2. However, both the proton and the electron are spin one-half particles. According to Chapter 6, when two spin one-half particles are combined (in the absence of orbital angular momentum) the net state has either spin 1 or spin 0. In fact, there are three spin 1 states, known as triplet states, and a single spin 0 state, known as the singlet state. For all states, the eigenvalues of $S_e^{\,2}$ and $S_p^{\,2}$ are $(3/4)\,\hbar^2$. The eigenvalue of S^2 is 0 for the singlet state, and $2\,\hbar^2$ for the triplet states. Hence,

$$\langle \mathbf{S}_p\cdot\mathbf{S}_e\rangle = -\frac{3}{4}\,\hbar^2 \tag{7.153}$$

for the singlet state, and

$$\langle \mathbf{S}_p\cdot\mathbf{S}_e\rangle = \frac{1}{4}\,\hbar^2 \tag{7.154}$$

for the triplet states.

It follows, from the above analysis, that spin-spin coupling breaks the degeneracy of the two $(1s)_{1/2}$ states of the hydrogen atom, lifting the energy of the triplet configuration, and lowering that of the singlet. This splitting is known as *hyperfine structure*. The net energy difference between the singlet and the triplet states is

$$\varDelta E_{000} = \frac{8}{3}\,g_p\,\frac{m_e}{m_p}\,\alpha^2\,|E_0| = 5.88\times 10^{-6}\,\mathrm{eV}, \tag{7.155}$$

where $|E_0| = 13.6\,\mathrm{eV}$ is the (magnitude of the) ground-state energy. Note that the hyperfine energy-shift is much smaller, by a factor m_e/m_p, than a typical fine structure energy-shift (see Exercise 7.4). If we convert the above energy into a wavelength then we obtain

$$\lambda = 21.1\,\mathrm{cm}. \tag{7.156}$$

This is the wavelength of the radiation emitted by a hydrogen atom that is collisionally excited from the singlet to the triplet state, and then decays back to the lower energy singlet state. The 21 cm line is famous in radio astronomy because it was used to map out the spiral structure of our galaxy in the 1950's.

Exercises

7.1 Calculate the energy-shift in the ground state of the one-dimensional harmonic oscillator when the perturbation

$$V = \lambda\,x^4$$

is added to

$$H = \frac{p_x^{\,2}}{2\,m} + \frac{1}{2}\,m\,\omega^2\,x^2.$$

The properly normalized ground-state wavefunction is

$$\psi(x) = \left(\frac{m\,\omega}{\pi\,\hbar}\right)^{1/4}\,\exp\left(-\frac{m\,\omega^2\,x^2}{2\,\hbar}\right).$$

7.2 Calculate the energy-shifts due to the first-order Stark effect in the $n = 3$ state of a hydrogen atom. You do not need to perform all of the integrals, but you should construct the correct linear combinations of states.

7.3 The Hamiltonian of the valence electron in a hydrogen-like atom can be written

$$H = \frac{p^2}{2\,m_e} + V(r) - \frac{p^4}{8\,m_e^3\,c^2}.$$

Here, the final term on the right-hand side is the first-order correction due to the electron's relativistic mass increase. Treating this term as a small perturbation, deduce that it causes an energy-shift in the energy eigenstate characterized by the standard quantum numbers n, l, m of

$$\Delta E_{nlm} = -\frac{1}{2\,m_e\,c^2}\left(E_n^{\,2} - 2\,E_n\,\langle V\rangle + \langle V^{\,2}\rangle\right),$$

where E_n is the unperturbed energy, and α the fine structure constant.

7.4 Consider an energy eigenstate of the hydrogen atom characterized by the standard quantum numbers n, l, and m. Show that if the energy-shift due to spin-orbit coupling (see Section 7.7) is added to that due to the electron's relativistic mass increase (see previous exercise) then the net fine structure energy-shift can be written

$$\Delta E_{nlm} = \frac{\alpha^2\,E_n}{n^2}\left(\frac{n}{j+1/2} - \frac{3}{4}\right).$$

Here, E_n is the unperturbed energy, α the fine structure constant, and $j = l \pm 1/2$ the quantum number associated with the magnitude of the sum of the electron's orbital and spin angular momenta. You will need to use the following standard results for a hydrogen atom:

$$\left\langle\frac{a_0}{r}\right\rangle = \frac{1}{n^2},$$

$$\left\langle\frac{a_0^2}{r^2}\right\rangle = \frac{1}{(l+1/2)\,n^3},$$

$$\left\langle\frac{a_0^3}{r^3}\right\rangle = \frac{1}{l\,(l+1/2)\,(l+1)\,n^3}.$$

Here, a_0 is the Bohr radius. Assuming that the above formula for the energy shift is valid for $l = 0$ (which it is), show that fine structure causes the energy of the $(2p)_{3/2}$ states of a hydrogen atom to exceed those of the $(2p)_{1/2}$ and $(2s)_{1/2}$ states by 4.5×10^{-5} eV.

8 Time-Dependent Perturbation Theory

8.1 Introduction

Suppose that the Hamiltonian of the system under consideration can be written

$$H = H_0 + H_1(t), \tag{8.1}$$

where H_0 does not contain time explicitly, and H_1 is a small time-dependent perturbation. It is assumed that we are able to calculate the eigenkets of the unperturbed Hamiltonian:

$$H_0\,|n\rangle = E_n\,|n\rangle. \tag{8.2}$$

We know that if the system is in one of the eigenstates of H_0 then, in the absence of the external perturbation, it remains in this state for ever. However, the presence of a small time-dependent perturbation can, in principle, give rise to a finite probability that a system initially in some eigenstate $|i\rangle$ of the unperturbed Hamiltonian is found in some other eigenstate at a subsequent time (because $|i\rangle$ is no longer an exact eigenstate of the total Hamiltonian). In other words, a time-dependent perturbation allows the system to make transitions between its unperturbed energy eigenstates. Let us investigate such transitions.

8.2 General Analysis

Suppose that at $t = t_0$ the state of the system is represented by

$$|A\rangle = \sum_n c_n\,|n\rangle, \tag{8.3}$$

where the c_n are complex numbers. Thus, the initial state is some linear superposition of the unperturbed energy eigenstates. In the absence of the time-dependent perturbation, the time evolution of the system is given by

$$|A, t_0, t\rangle = \sum_n c_n \exp[-{\rm i}\,E_n\,(t-t_0)/\hbar]\,|n\rangle. \tag{8.4}$$

Now, the probability of finding the system in state $|n\rangle$ at time t is

$$P_n(t) = |c_n \exp[-{\rm i}\,E_n(t-t_0)/\hbar]|^2 = |c_n|^2 = P_n(t_0). \tag{8.5}$$

Clearly, with $H_1 = 0$, the probability of finding the system in state $|n\rangle$ at time t is exactly the same as the probability of finding the system in this state at the initial time t_0. However, with $H_1 \neq 0$, we expect $P_n(t)$ to vary with time. Thus, we can write

$$|A, t_0, t\rangle = \sum_n c_n(t) \exp[-{\rm i}\,E_n\,(t-t_0)/\hbar]\,|n\rangle, \tag{8.6}$$

where $P_n(t) = |c_n(t)|^2$. Here, we have carefully separated the fast phase oscillation of the eigenkets, which depends on the unperturbed Hamiltonian, from the slow variation of the amplitudes $c_n(t)$, which depends entirely on the perturbation (i.e., c_n is constant if $H_1 = 0$). Note that the eigenkets $|n\rangle$, appearing in Equation (8.6), are *time-independent* (they are actually the eigenkets of H_0 evaluated at the time t_0).

Schrödinger's time evolution equation yields

$$\mathrm{i}\,\hbar\,\frac{\partial}{\partial t}\,|A,t_0,t\rangle = H\,|A,t_0,t\rangle = (H_0+H_1)\,|A,t_0,t\rangle. \tag{8.7}$$

It follows from Equation (8.6) that

$$(H_0+H_1)\,|A,t_0,t\rangle = \sum_m c_m(t)\exp[-\mathrm{i}\,E_m\,(t-t_0)/\hbar]\,(E_m+H_1)\,|m\rangle. \tag{8.8}$$

We also have

$$\mathrm{i}\,\hbar\,\frac{\partial}{\partial t}\,|A,t_0,t\rangle = \sum_m\left(\mathrm{i}\,\hbar\,\frac{dc_m}{dt}+c_m(t)\,E_m\right)\exp[-\mathrm{i}\,E_m\,(t-t_0)/\hbar]\,|m\rangle, \tag{8.9}$$

where use has been made of the time-independence of the kets $|m\rangle$. According to Equation (8.7), we can equate the right-hand sides of the previous two equations to obtain

$$\sum_m \mathrm{i}\,\hbar\,\frac{dc_m}{dt}\exp[-\mathrm{i}\,E_m\,(t-t_0)/\hbar]\,|m\rangle = \sum_m c_m(t)\exp[-\mathrm{i}\,E_m\,(t-t_0)/\hbar]\,H_1\,|m\rangle. \tag{8.10}$$

Left-multiplication by $\langle n|$ yields

$$\mathrm{i}\,\hbar\,\frac{dc_n}{dt} = \sum_m H_{nm}(t)\,\exp[\,\mathrm{i}\,\omega_{nm}\,(t-t_0)]\,c_m(t), \tag{8.11}$$

where

$$H_{nm}(t) = \langle n|\,H_1(t)\,|m\rangle, \tag{8.12}$$

and

$$\omega_{nm} = \frac{E_n-E_m}{\hbar}. \tag{8.13}$$

Here, we have made use of the standard orthonormality result, $\langle n|m\rangle = \delta_{nm}$. Suppose that there are N linearly independent eigenkets of the unperturbed Hamiltonian. According to Equation (8.11), the time variation of the coefficients c_n, which specify the probability of finding the system in state $|n\rangle$ at time t, is determined by N coupled first-order differential equations. Note that Equation (8.11) is exact—we have made no approximations at this stage. Unfortunately, we cannot generally find exact solutions to this equation, so we have to obtain approximate solutions via suitable expansions in small quantities. However, for the particularly simple case of a two-state system (i.e., $N = 2$), it is actually possible to solve Equation (8.11) without approximation. This solution is of great practical importance.

8.3 Two-State System

Consider a system in which the time-independent Hamiltonian possesses two eigenstates, denoted

$$H_0\,|1\rangle = E_1\,|1\rangle, \tag{8.14}$$

$$H_0\,|2\rangle = E_2\,|2\rangle. \tag{8.15}$$

Suppose, for the sake of simplicity, that the diagonal matrix elements of the interaction Hamiltonian, H_1, are zero:

$$\langle 1|\,H_1\,|1\rangle = \langle 2|\,H_1\,|2\rangle = 0. \tag{8.16}$$

The off-diagonal matrix elements are assumed to oscillate sinusoidally at some frequency ω:

$$\langle 1|\,H_1\,|2\rangle = \langle 2|\,H_1\,|1\rangle^{*} = \gamma\exp(\,\mathrm{i}\,\omega\,t), \tag{8.17}$$

where γ and ω are real. Note that it is only the off-diagonal matrix elements that give rise to the effect which we are interested in—namely, transitions between states 1 and 2.

For a two-state system, Equation (8.11) reduces to

$$\mathrm{i}\,\hbar\,\frac{dc_1}{dt} = \gamma\exp[+\mathrm{i}\,(\omega-\omega_{21})\,t\,]\,c_2, \tag{8.18}$$

$$\mathrm{i}\,\hbar\,\frac{dc_2}{dt} = \gamma\exp[-\mathrm{i}\,(\omega-\omega_{21})\,t\,]\,c_1, \tag{8.19}$$

where $\omega_{21} = (E_2 - E_1)/\hbar$, and it is assumed that $t_0 = 0$. Equations (8.18) and (8.19) can be combined to give a second-order differential equation for the time variation of the amplitude c_2:

$$\frac{d^2c_2}{dt^2} + \mathrm{i}\,(\omega-\omega_{21})\,\frac{dc_2}{dt} + \frac{\gamma^2}{\hbar^2}\,c_2 = 0. \tag{8.20}$$

Once we have solved for c_2, we can use Equation (8.19) to obtain the amplitude c_1. Let us look for a solution in which the system is certain to be in state 1 at time $t = 0$. Thus, our initial conditions are $c_1(0) = 1$ and $c_2(0) = 0$. It is easily demonstrated that the appropriate solutions are

$$c_2(t) = \frac{-\mathrm{i}\,\gamma/\hbar}{[\gamma^2/\hbar^2 + (\omega-\omega_{21})^2/4]^{1/2}}\,\exp[-\mathrm{i}\,(\omega-\omega_{21})\,t/2]\,\sin\Big([\gamma^2/\hbar^2 + (\omega-\omega_{21})^2/4]^{1/2}\,t\Big), \tag{8.21}$$

$$\begin{aligned} c_1(t) = {} & \exp[\,\mathrm{i}\,(\omega-\omega_{21})\,t/2]\,\cos\Big([\gamma^2/\hbar^2 + (\omega-\omega_{21})^2/4]^{1/2}\,t\Big) \\ & - \frac{\mathrm{i}\,(\omega-\omega_{21})/2}{[\gamma^2/\hbar^2 + (\omega-\omega_{21})^2/4]^{1/2}}\,\exp[\,\mathrm{i}\,(\omega-\omega_{21})\,t/2]\,\sin\Big([\gamma^2/\hbar^2 + (\omega-\omega_{21})^2/4]^{1/2}\,t\Big). \end{aligned} \tag{8.22}$$

The probability of finding the system in state 1 at time t is simply $P_1(t) = |c_1|^2$. Likewise, the probability of finding the system in state 2 at time t is $P_2(t) = |c_2|^2$. It follows that

$$P_2(t) = \frac{\gamma^2/\hbar^2}{\gamma^2/\hbar^2 + (\omega-\omega_{21})^2/4}\,\sin^2\Big([\gamma^2/\hbar^2 + (\omega-\omega_{21})^2/4]^{1/2}\,t\Big), \tag{8.23}$$

$$P_1(t) = 1 - P_2(t). \tag{8.24}$$

Equation (8.23) exhibits all the features of a classic *resonance*. At resonance, when the oscillation frequency of the perturbation, ω, matches the frequency ω_{21}, we find that

$$P_1(t) = \cos^2(\gamma\,t/\hbar), \tag{8.25}$$

$$P_2(t) = \sin^2(\gamma\,t/\hbar). \tag{8.26}$$

According to the above result, the system starts off at $t = 0$ in state 1. After a time interval $\pi\,\hbar/2\,\gamma$, it is certain to be in state 2. After a further time interval $\pi\,\hbar/2\,\gamma$, it is certain to be in state 1, and so on. In other words, the system periodically flip-flops between states 1 and 2 under the influence of the time-dependent perturbation. This implies that the system alternatively absorbs and emits energy from the source of the perturbation.

The absorption-emission cycle also take place away from the resonance, when $\omega \neq \omega_{21}$. However, the amplitude of oscillation of the coefficient c_2 is reduced. This means that the maximum value of $P_2(t)$ is no longer unity, nor is the minimum value of $P_1(t)$ zero. In fact, if we plot the maximum value of $P_2(t)$ as a function of the applied frequency, ω, then we obtain a resonance curve whose maximum (unity) lies at the resonance, and whose full-width half-maximum (in frequency) is $4\,\gamma/\hbar$. Thus, if the applied frequency differs from the resonant frequency by substantially more than $2\,\gamma/\hbar$ then the probability of the system jumping from state 1 to state 2 is very small. In other words, the time-dependent perturbation is only effective at causing transitions between states 1 and 2 if its frequency of oscillation lies in the approximate range $\omega_{21} \pm 2\,\gamma/\hbar$. Clearly, the weaker the perturbation (i.e., the smaller γ becomes), the narrower the resonance.

8.4 Spin Magnetic Resonance

Consider a bound electron placed in a uniform z-directed magnetic field, and then subjected to a small time-dependent magnetic field rotating in the x-y plane. Thus,

$$\mathbf{B} = B_0\,\mathbf{e}_z + B_1\left[\cos(\omega\,t)\,\mathbf{e}_x + \sin(\omega\,t)\,\mathbf{e}_y\right], \tag{8.27}$$

where B_0 and B_1 are constants, with $B_1 \ll B_0$. The rotating magnetic field usually represents the magnetic component of an electromagnetic wave propagating along the z-axis. In this system, the electric component of the wave has no effect. The Hamiltonian is written

$$H = -\boldsymbol{\mu}\cdot\mathbf{B} = H_0 + H_1, \tag{8.28}$$

where

$$H_0 = \frac{e\,B_0}{m_e}\,S_z, \tag{8.29}$$

and

$$H_1 = \frac{e\,B_1}{m_e}\left[\cos(\omega\,t)\,S_x + \sin(\omega\,t)\,S_y\right]. \tag{8.30}$$

The eigenstates of the unperturbed Hamiltonian are the 'spin up' and 'spin down' states, denoted $|+\rangle$ and $|-\rangle$, respectively. Thus,

$$H_0\,|\pm\rangle = \pm\frac{e\,\hbar\,B_0}{2\,m_e}\,|\pm\rangle. \tag{8.31}$$

The time-dependent Hamiltonian can be written

$$H_1 = \frac{e\,B_1}{2\,m_e}\,[\exp(\,\mathrm{i}\,\omega\,t)\,S^- + \exp(-\mathrm{i}\,\omega\,t)\,S^+]\,, \tag{8.32}$$

where $S^\pm = S_x \pm \mathrm{i}\,S_y$ are the conventional raising and lowering operators for spin angular momentum. It follows that

$$\langle +|\,H_1\,|+\rangle = \langle -|\,H_1\,|-\rangle = 0, \tag{8.33}$$

and

$$\langle -|\,H_1\,|+\rangle = \langle +|\,H_1\,|-\rangle^* = \frac{e\,\hbar\,B_1}{2\,m_e}\exp(\,\mathrm{i}\,\omega\,t). \tag{8.34}$$

It can be seen that this system is exactly the same as the two-state system discussed in the previous section, provided that we make the identifications

$$|1\rangle \rightarrow |-\rangle, \tag{8.35}$$

$$|2\rangle \rightarrow |+\rangle, \tag{8.36}$$

$$\omega_{21} \rightarrow \frac{e\,B_0}{m_e}, \tag{8.37}$$

$$\gamma \rightarrow \frac{e\,\hbar\,B_1}{2\,m_e}. \tag{8.38}$$

The resonant frequency, ω_{21}, is simply the spin precession frequency for an electron in a uniform magnetic field of strength B_0. In the absence of the perturbation, the expectation values of S_x and S_y oscillate because of the spin precession, but the expectation value of S_z remains invariant. If we now apply a magnetic perturbation rotating at the resonant frequency then, according to the analysis of the previous section, the system undergoes a succession of spin-flops, $|+\rangle \rightleftharpoons |-\rangle$, in addition to the spin precession. We also know that if the oscillation frequency of the applied field is very different from the resonant frequency then there is virtually zero probability of the field triggering a spin-flop. The width of the resonance (in frequency) is determined by the strength of the oscillating magnetic perturbation. Experimentalist are able to measure the magnetic moments of electrons, and other spin one-half particles, to a high degree of accuracy by placing the particles in a magnetic field, and subjecting them to an oscillating magnetic field whose frequency is gradually scanned. By determining the resonant frequency (i.e., the frequency at which the particles absorb energy from the oscillating field), it is possible to calculate the magnetic moment.

8.5 Dyson Series

Let us now try to find approximate solutions of Equation (8.11) for a general system. It is convenient to work in terms of the time evolution operator, $U(t_0, t)$, which is defined

$$|A, t_0, t\rangle = U(t_0, t)\,|A\rangle. \tag{8.39}$$

Here, $|A, t_0, t\rangle$ is the state ket of the system at time t, given that the state ket at the initial time t_0 is $|A\rangle$. It is easily seen that the time evolution operator satisfies the differential equation

$$\mathrm{i}\,\hbar\,\frac{\partial U(t_0, t)}{\partial t} = (H_0 + H_1)\,U(t_0, t), \tag{8.40}$$

subject to the initial condition

$$U(t_0, t_0) = 1. \tag{8.41}$$

In the absence of the external perturbation, the time evolution operator reduces to

$$U(t_0, t) = \exp[-\mathrm{i}\,H_0\,(t - t_0)/\hbar]. \tag{8.42}$$

Let us switch on the perturbation and look for a solution of the form

$$U(t_0, t) = \exp[-\mathrm{i}\,H_0\,(t - t_0)/\hbar]\,U_I(t_0, t). \tag{8.43}$$

It is readily demonstrated that U_I satisfies the differential equation

$$\mathrm{i}\,\hbar\,\frac{\partial U_I(t_0, t)}{\partial t} = H_I(t_0, t)\,U_I(t_0, t), \tag{8.44}$$

where

$$H_I(t_0, t) = \exp[+\mathrm{i}\,H_0\,(t - t_0)/\hbar]\,H_1\,\exp[-\mathrm{i}\,H_0\,(t - t_0)/\hbar], \tag{8.45}$$

subject to the initial condition

$$U_I(t_0, t_0) = 1. \tag{8.46}$$

Note that U_I specifies that component of the time evolution operator which is due to the time-dependent perturbation. Thus, we would expect U_I to contain all of the information regarding transitions between different eigenstates of H_0 caused by the perturbation.

Suppose that the system starts off at time t_0 in the eigenstate $|i\rangle$ of the unperturbed Hamiltonian. The subsequent evolution of the state ket is given by Equation (8.6),

$$|i, t_0, t\rangle = \sum_m c_m(t)\exp[-\mathrm{i}\,E_m\,(t - t_0)/\hbar]\,|m\rangle. \tag{8.47}$$

However, we also have

$$|i, t_0, t\rangle = \exp[-\mathrm{i}\,H_0\,(t - t_0)/\hbar]\,U_I(t_0, t)\,|i\rangle. \tag{8.48}$$

It follows that

$$c_n(t) = \langle n|\,U_I(t_0, t)\,|i\rangle, \tag{8.49}$$

where use has been made of $\langle n|m\rangle = \delta_{n\,m}$. Thus, the probability that the system is found in state $|n\rangle$ at time t, given that it is definitely in state $|i\rangle$ at time t_0, is simply

$$P_{i\rightarrow n}(t_0, t) = |\langle n|\,U_I(t_0, t)\,|i\rangle|^2. \tag{8.50}$$

This quantity is usually termed the *transition probability* between states $|i\rangle$ and $|n\rangle$.

Note that the differential equation (8.44), plus the initial condition (8.46), are equivalent to the following integral equation,

$$U_I(t_0,t) = 1 - \frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} dt'\, H_I(t_0,t')\, U_I(t_0,t'). \tag{8.51}$$

We can obtain an approximate solution to this equation by iteration:

$$\begin{aligned} U_I(t_0,t) &\simeq 1 - \frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} H_I(t_0,t')\left[1 - \frac{\mathrm{i}}{\hbar}\int_{t_0}^{t'} dt'\, H_I(t_0,t'')\, U_I(t_0,t'')\right] \\ &\simeq 1 - \frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} H_I(t_0,t')\, dt' + \left(\frac{-\mathrm{i}}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\, H_I(t_0,t')\, H_I(t_0,t'') + \cdots\,. \end{aligned} \tag{8.52}$$

This expansion is known as the *Dyson series*. Let

$$c_n = c_n^{(0)} + c_n^{(1)} + c_n^{(2)} + \cdots\,, \tag{8.53}$$

where the superscript $^{(1)}$ refers to a first-order term in the expansion, etc. It follows from Equations (8.49) and (8.52) that

$$c_n^{(0)}(t) = \delta_{i\,n}, \tag{8.54}$$

$$c_n^{(1)}(t) = -\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} dt'\, \langle n|\, H_I(t_0,t')\, |i\rangle, \tag{8.55}$$

$$c_n^{(2)}(t) = \left(\frac{-\mathrm{i}}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\, \langle n|\, H_I(t_0,t')\, H_I(t_0,t'')\, |i\rangle. \tag{8.56}$$

These expressions simplify to

$$c_n^{(0)}(t) = \delta_{in}, \tag{8.57}$$

$$c_n^{(1)}(t) = -\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} dt'\; \exp[\,\mathrm{i}\,\omega_{ni}\,(t'-t_0)]\, H_{ni}(t'), \tag{8.58}$$

$$c_n^{(2)}(t) = \left(\frac{-\mathrm{i}}{\hbar}\right)^2 \sum_m \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\; \exp[\,\mathrm{i}\,\omega_{nm}\,(t'-t_0)]\, H_{nm}(t')\; \exp[\,\mathrm{i}\,\omega_{mi}\,(t''-t_0)]\, H_{mi}(t''), \tag{8.59}$$

where

$$\omega_{nm} = \frac{E_n - E_m}{\hbar}, \tag{8.60}$$

and

$$H_{nm}(t) = \langle n|\, H_1(t)\, |m\rangle. \tag{8.61}$$

The transition probability between states *i* and *n* is simply

$$P_{i\to n}(t_0,t) = |c_n^{(0)} + c_n^{(1)} + c_n^{(2)} + \cdots|^2. \tag{8.62}$$

According to the above analysis, there is no chance of a transition between states $|i\rangle$ and $|n\rangle$ (where $i \neq n$) to zeroth order (i.e., in the absence of the perturbation). To first order, the transition probability is proportional to the time integral of the matrix element $\langle n|\,H_1\,|i\rangle$, weighted by some oscillatory phase-factor. Thus, if the matrix element is zero then there is no chance of a first-order transition between states $|i\rangle$ and $|n\rangle$. However, to second order, a transition between states $|i\rangle$ and $|n\rangle$ is possible even when the matrix element $\langle n|\,H_1\,|i\rangle$ is zero.

8.6 Sudden Perturbations

Consider, for example, a constant perturbation that is suddenly switched on at time $t = 0$:

$$\begin{aligned} H_1(t) &= 0 && \text{for } t < 0\\ H_1(t) &= H_1 && \text{for } t \geq 0, \end{aligned} \tag{8.63}$$

where H_1 is time-independent, but is generally a function of the position, momentum, and spin operators. Suppose that the system is definitely in state $|i\rangle$ at time $t = 0$. According to Equations (8.57)–(8.59) (with $t_0 = 0$),

$$c_n^{(0)}(t) = \delta_{i\,n}, \tag{8.64}$$

$$c_n^{(1)}(t) = -\frac{\mathrm{i}}{\hbar}\,H_{ni}\int_0^t dt'\,\exp[\,\mathrm{i}\,\omega_{ni}\,(t'-t)] = \frac{H_{ni}}{E_n - E_i}\,[1 - \exp(\,\mathrm{i}\,\omega_{ni}\,t)], \tag{8.65}$$

giving

$$P_{i\to n}(t) \simeq |c_n^{(1)}|^2 = \frac{4\,|H_{ni}|^2}{|E_n - E_i|^2}\,\sin^2\left[\frac{(E_n - E_i)\,t}{2\,\hbar}\right], \tag{8.66}$$

for $i \neq n$. The transition probability between states $|i\rangle$ and $|n\rangle$ can be written

$$P_{i\to n}(t) = \frac{|H_{ni}|^2\,t^2}{\hbar^2}\,\mathrm{sinc}^2\left[\frac{(E_n - E_i)\,t}{2\,\hbar}\right], \tag{8.67}$$

where

$$\mathrm{sinc}(x) \equiv \frac{\sin x}{x}. \tag{8.68}$$

The sinc function is highly oscillatory, and decays like $1/|x|$ at large $|x|$. It is a good approximation to say that $\mathrm{sinc}(x)$ is small except when $|x| \lesssim \pi$. It follows that the transition probability, $P_{i\to n}$, is small except when

$$|E_n - E_i| \lesssim \frac{2\pi\,\hbar}{t}. \tag{8.69}$$

Note that in the limit $t \to \infty$ only those transitions that conserve energy (i.e., $E_n = E_i$) have an appreciable probability of occurrence. At finite t, is is possible to have transitions which do not exactly conserve energy, provided that

$$\Delta E\,\Delta t \lesssim h, \tag{8.70}$$

where $\Delta E = |E_n - E_i|$ is the change in energy of the system associated with the transition, and $\Delta t = t$ is the time elapsed since the perturbation was switched on. This result is just a manifestation of the well-known uncertainty relation for energy and time. Incidentally, the energy-time uncertainty relation is fundamentally different to the position-momentum uncertainty relation, because (in non-relativistic quantum mechanics) position and momentum are operators, whereas time is merely a parameter.

The probability of a transition that conserves energy (i.e., $E_n = E_i$) is

$$P_{i\rightarrow n}(t) = \frac{|H_{in}|^2\,t^2}{\hbar^2}, \tag{8.71}$$

where use has been made of sinc(0) = 1. Note that this probability grows *quadratically* with time. This result is somewhat surprising, because it implies that the probability of a transition occurring in a fixed time interval, t to $t + dt$, grows linearly with t, despite the fact that H_1 is constant for $t > 0$. In practice, there is usually a group of final states, all possessing nearly the same energy as the energy of the initial state $|i\rangle$. It is helpful to define the density of states, $\rho(E)$, where the number of final states lying in the energy range E to $E + dE$ is given by $\rho(E)\,dE$. Thus, the probability of a transition from the initial state i to any of the continuum of possible final states is

$$P_{i\rightarrow}(t) = \int dE_n\, P_{i\rightarrow n}(t)\,\rho(E_n), \tag{8.72}$$

giving

$$P_{i\rightarrow}(t) = \frac{2\,t}{\hbar}\int dx\,|H_{ni}|^2\,\rho(E_n)\,\mathrm{sinc}^2(x), \tag{8.73}$$

where

$$x = (E_n - E_i)\,t/2\,\hbar, \tag{8.74}$$

and use has been made of Equation (8.67). We know that in the limit $t \rightarrow \infty$ the function sinc(x) is only non-zero in an infinitesimally narrow range of final energies centered on $E_n = E_i$. It follows that, in this limit, we can take $\rho(E_n)$ and $|H_{ni}|^2$ out of the integral in the above formula to obtain

$$P_{i\rightarrow[n]}(t) = \frac{2\pi}{\hbar}\,\overline{|H_{ni}|^2}\,\rho(E_n)\,t\,\Big|_{E_n\simeq E_i}, \tag{8.75}$$

where $P_{i\rightarrow[n]}$ denotes the transition probability between the initial state $|i\rangle$ and all final states $|n\rangle$ that have approximately the same energy as the initial state. Here, $\overline{|H_{ni}|^2}$ is the average of $|H_{ni}|^2$ over all final states with approximately the same energy as the initial state. In deriving the above formula, we have made use of the result

$$\int_{-\infty}^{\infty} dx\;\mathrm{sinc}^2(x) = \pi. \tag{8.76}$$

Note that the transition probability, $P_{i\rightarrow[n]}$, is now proportional to t, instead of t^2.

It is convenient to define the *transition rate*, which is simply the transition probability per unit time. Thus,

$$w_{i\rightarrow[n]} = \frac{dP_{i\rightarrow[n]}}{dt}, \tag{8.77}$$

giving

$$w_{i\to[n]} = \frac{2\pi}{\hbar}\,\overline{|H_{ni}|^2}\,\rho(E_n)\bigg|_{E_n\simeq E_i}. \tag{8.78}$$

This appealingly simple result is known as *Fermi's golden rule*. Note that the transition rate is constant in time (for $t > 0$): i.e., the probability of a transition occurring in the time interval t to $t + dt$ is independent of t for fixed dt. Fermi's golden rule is sometimes written

$$w_{i\to n} = \frac{2\pi}{\hbar}\,|H_{ni}|^2\,\delta(E_n - E), \tag{8.79}$$

where it is understood that this formula must be integrated with $\int dE_n\,\rho(E_n)$ to obtain the actual transition rate.

Let us now calculate the second-order term in the Dyson series, using the constant perturbation (8.63). From Equation (8.59) we find that

$$\begin{aligned}
c_n^{(2)}(t) &= \left(\frac{-\mathrm{i}}{\hbar}\right)^2 \sum_m H_{nm}H_{mi}\int_0^t dt'\,\exp(\,\mathrm{i}\,\omega_{nm}\,t'\,)\int_0^{t'} dt''\,\exp(\,\mathrm{i}\,\omega_{mi}\,t\,)\\
&= \frac{\mathrm{i}}{\hbar}\sum_m \frac{H_{nm}\,H_{mi}}{E_m - E_i}\int_0^t dt'\,[\exp(\,\mathrm{i}\,\omega_{ni}\,t'\,) - \exp(\,\mathrm{i}\,\omega_{nm}\,t']\,)\\
&= \frac{\mathrm{i}\,t}{\hbar}\sum_m \frac{H_{nm}H_{mi}}{E_m - E_i}\,[\exp(\,\mathrm{i}\,\omega_{ni}\,t/2)\,\mathrm{sinc}(\omega_{ni}\,t/2) - \exp(\,\mathrm{i}\,\omega_{nm}\,t/2)\,\mathrm{sinc}(\omega_{nm}\,t/2)]\,.
\end{aligned} \tag{8.80}$$

Thus,

$$\begin{aligned}
c_n(t) = c_n^{(1)} + c_n^{(2)} = \frac{\mathrm{i}\,t}{\hbar}\exp(\,\mathrm{i}\,\omega_{ni}\,t/2)\,\Bigg[&\left(H_{ni} + \sum_m \frac{H_{nm}\,H_{mi}}{E_m - E_i}\right)\mathrm{sinc}(\omega_{ni}\,t/2)\\
&- \sum_m \frac{H_{nm}\,H_{mi}}{E_m - E_i}\exp(\,\mathrm{i}\,\omega_{im}\,t/2)\,\mathrm{sinc}(\omega_{nm}\,t/2)\Bigg],
\end{aligned} \tag{8.81}$$

where use has been made of Equation (8.65). It follows, by analogy with the previous analysis, that

$$w_{i\to[n]} = \frac{2\pi}{\hbar}\,\overline{\left|H_{ni} + \sum_m \frac{H_{nm}\,H_{mi}}{E_m - E_i}\right|^2}\,\rho(E_n)\Bigg|_{E_n\simeq E_i}, \tag{8.82}$$

where the transition rate is calculated for all final states, $|n\rangle$, with approximately the same energy as the initial state, $|i\rangle$, and for intermediate states, $|m\rangle$ whose energies differ from that of the initial state. The fact that $E_m \neq E_i$ causes the last term on the right-hand side of Equation (8.81) to average to zero (due to the oscillatory phase-factor) during the evaluation of the transition probability.

According to Equation (8.82), a second-order transition takes place in two steps. First, the system makes a non-energy-conserving transition to some intermediate state $|m\rangle$. Subsequently, the system makes another non-energy-conserving transition to the final state $|n\rangle$. The net transition, from $|i\rangle$ to $|n\rangle$, conserves energy. The non-energy-conserving transitions are generally termed *virtual transitions*, whereas the energy conserving first-order transition is termed a *real transition*. The above formula clearly breaks down if $H_{nm}\,H_{mi} \neq 0$ when $E_m = E_i$. This problem can be avoided by gradually turning on the perturbation: i.e., $H_1 \rightarrow \exp(\eta\,t)\,H_1$ (where η is very small). The net result is to change the energy denominator in Equation (8.82) from $E_i - E_m$ to $E_i - E_m + {\rm i}\,\hbar\,\eta$.

8.7 Energy-Shifts and Decay-Widths

We have examined how a state $|n\rangle$, other than the initial state $|i\rangle$, becomes populated as a result of some time-dependent perturbation applied to the system. Let us now consider how the initial state becomes depopulated.

In this case, it is convenient to gradually turn on the perturbation from zero at $t = -\infty$. Thus,

$$H_1(t) = \exp(\eta\,t)\,H_1, \tag{8.83}$$

where η is small and positive, and H_1 is a constant.

In the remote past, $t \rightarrow -\infty$, the system is assumed to be in the initial state $|i\rangle$. Thus, $c_i(t \rightarrow -\infty) = 1$, and $c_{n\neq i}(t \rightarrow -\infty) = 0$. Basically, we want to calculate the time evolution of the coefficient $c_i(t)$. First, however, let us check that our previous Fermi golden rule result still applies when the perturbing potential is turned on slowly, instead of very suddenly. For $c_{n\neq i}(t)$ we have from Equations (8.57)–(8.58) that

$$c_n^{(0)}(t) = 0, \tag{8.84}$$

$$c_n^{(1)}(t) = -\frac{\rm i}{\hbar}\,H_{ni}\int_{-\infty}^{t} dt'\,\exp[(\eta + {\rm i}\,\omega_{ni})\,t'] = -\frac{\rm i}{\hbar}\,H_{ni}\,\frac{\exp[(\eta + {\rm i}\,\omega_{ni})\,t]}{\eta + {\rm i}\,\omega_{ni}}, \tag{8.85}$$

where $H_{ni} = \langle n|\,H_1\,|i\rangle$. It follows that, to first order, the transition probability from state $|i\rangle$ to state $|n\rangle$ is

$$P_{i\rightarrow n}(t) = |c_n^{(1)}|^2 = \frac{|H_{ni}|^2}{\hbar^2}\,\frac{\exp(2\,\eta\,t)}{\eta^2 + \omega_{ni}^2}. \tag{8.86}$$

The transition rate is given by

$$w_{i\rightarrow n}(t) = \frac{dP_{i\rightarrow n}}{dt} = \frac{2\,|H_{ni}|^2}{\hbar^2}\,\frac{\eta\,\exp(2\,\eta\,t)}{\eta^2 + \omega_{ni}^2}. \tag{8.87}$$

Consider the limit $\eta \rightarrow 0$. In this limit, $\exp(\eta\,t) \rightarrow 1$, but

$$\lim_{\eta\rightarrow 0}\frac{\eta}{\eta^2 + \omega_{ni}^2} = \pi\,\delta(\omega_{ni}) = \pi\,\hbar\,\delta(E_n - E_i). \tag{8.88}$$

Thus, Equation (8.87) yields the standard Fermi golden rule result

$$w_{i\rightarrow n} = \frac{2\pi}{\hbar}\,|H_{ni}|^2\,\delta(E_n - E_i). \tag{8.89}$$

It is clear that the delta-function in the above formula actually represents a function that is highly peaked at some particular energy. The width of the peak is determined by how fast the perturbation is switched on.

Let us now calculate $c_i(t)$ using Equations (8.57)–(8.59). We have

$$c_i^{(0)}(t) = 1, \tag{8.90}$$

$$c_i^{(1)}(t) = -\frac{\mathrm{i}}{\hbar}\,H_{ii}\int_{-\infty}^{t}\exp(\eta\,t')\,dt' = -\frac{\mathrm{i}}{\hbar}\,H_{ii}\,\frac{\exp(\eta\,t)}{\eta}, \tag{8.91}$$

$$\begin{aligned} c_i^{(2)}(t) &= \left(\frac{-\mathrm{i}}{\hbar}\right)^2\sum_m |H_{mi}|^2\int_{-\infty}^{t}dt'\int_{-\infty}^{t'}dt''\;\exp[(\eta+\mathrm{i}\,\omega_{im})\,t']\exp[(\eta+\mathrm{i}\,\omega_{mi})\,t''] \\ &= \left(\frac{-\mathrm{i}}{\hbar}\right)^2\sum_m |H_{mi}|^2\frac{\exp(2\,\eta\,t)}{2\,\eta\,(\eta+\mathrm{i}\,\omega_{mi})}. \end{aligned} \tag{8.92}$$

Thus, to second order we have

$$c_i(t) \simeq 1 + \left(\frac{-\mathrm{i}}{\hbar}\right)H_{ii}\,\frac{\exp(\eta\,t)}{\eta} + \left(\frac{-\mathrm{i}}{\hbar}\right)^2|H_{ii}|^2\,\frac{\exp(2\,\eta\,t)}{2\,\eta^2} + \left(\frac{-\mathrm{i}}{\hbar}\right)\sum_{m\neq i}\frac{|H_{mi}|^2\exp(2\,\eta\,t)}{2\,\eta\,(E_i - E_m + \mathrm{i}\,\hbar\,\eta)}. \tag{8.93}$$

Let us now consider the ratio $\dot{c}_i/c_i$, where $\dot{c}_i \equiv dc_i/dt$. Using Equation (8.93), we can evaluate this ratio in the limit $\eta \rightarrow 0$. We obtain

$$\begin{aligned} \frac{\dot{c}_i}{c_i} &\simeq \left[\left(\frac{-\mathrm{i}}{\hbar}\right)H_{ii} + \left(\frac{-\mathrm{i}}{\hbar}\right)^2\frac{|H_{ii}|^2}{\eta} + \left(\frac{-\mathrm{i}}{\hbar}\right)\sum_{m\neq i}\frac{|H_{mi}|^2}{E_i - E_m + \mathrm{i}\,\hbar\,\eta}\right]\Bigg/\left(1 - \frac{\mathrm{i}}{\hbar}\frac{H_{ii}}{\eta}\right) \\ &\simeq \left(\frac{-\mathrm{i}}{\hbar}\right)H_{ii} + \lim_{\eta\rightarrow 0}\left(\frac{-\mathrm{i}}{\hbar}\right)\sum_{m\neq i}\frac{|H_{mi}|^2}{E_i - E_m + \mathrm{i}\,\hbar\,\eta}. \end{aligned} \tag{8.94}$$

This result is formally correct to second order in perturbed quantities. Note that the right-hand side of Equation (8.94) is independent of time. We can write

$$\frac{\dot{c}_i}{c_i} = \left(\frac{-\mathrm{i}}{\hbar}\right)\mathit{\Delta}_i, \tag{8.95}$$

where

$$\mathit{\Delta}_i = H_{ii} + \lim_{\eta\rightarrow 0}\sum_{m\neq i}\frac{|H_{mi}|^2}{E_i - E_m + \mathrm{i}\,\hbar\,\eta} \tag{8.96}$$

is a constant. According to a well-known result in pure mathematics,

$$\lim_{\epsilon\rightarrow 0}\frac{1}{x + \mathrm{i}\,\epsilon} = \mathcal{P}\,\frac{1}{x} - \mathrm{i}\,\pi\,\delta(x), \tag{8.97}$$

where $\epsilon > 0$, and $\mathcal{P}$ denotes the principal part. It follows that

$$\mathit{\Delta}_i = H_{ii} + \mathcal{P}\sum_{m\neq i}\frac{|H_{mi}|^2}{E_i - E_m} - \mathrm{i}\,\pi\sum_{m\neq i}|H_{mi}|^2\,\delta(E_i - E_m). \tag{8.98}$$

It is convenient to normalize the solution of Equation (8.95) such that $c_i(0) = 1$. Thus, we obtain

$$c_i(t) = \exp\!\left(\frac{-\mathrm{i}\,\mathit{\Delta}_i\,t}{\hbar}\right). \tag{8.99}$$

According to Equation (8.6), the time evolution of the initial state ket $|i\rangle$ is given by

$$|i,t\rangle = \exp[-\mathrm{i}\,(\mathit{\Delta}_i + E_i)\,t/\hbar]\,|i\rangle. \tag{8.100}$$

We can rewrite this result as

$$|i,t\rangle = \exp(-\mathrm{i}\,[E_i + \mathrm{Re}(\mathit{\Delta}_i)\,]\,t/\hbar)\,\exp[\,\mathrm{Im}(\mathit{\Delta}_i)\,t/\hbar]\,|i\rangle. \tag{8.101}$$

It is clear that the real part of $\mathit{\Delta}_i$ gives rise to a simple shift in energy of state $|i\rangle$, whereas the imaginary part of $\mathit{\Delta}_i$ governs the growth or decay of this state. Thus,

$$|i,t\rangle = \exp[-\mathrm{i}\,(E_i + \mathit{\Delta}E_i)\,t/\hbar]\,\exp(-\mathit{\Gamma}_i\,t/2\,\hbar)\,|i\rangle, \tag{8.102}$$

where

$$\mathit{\Delta}E_i = \mathrm{Re}(\mathit{\Delta}_i) = H_{ii} + \mathcal{P}\sum_{m\neq i}\frac{|H_{mi}|^2}{E_i - E_m}, \tag{8.103}$$

and

$$\frac{\mathit{\Gamma}_i}{\hbar} = -\frac{2\,\mathrm{Im}(\mathit{\Delta}_i)}{\hbar} = \frac{2\pi}{\hbar}\sum_{m\neq i}|H_{mi}|^2\,\delta(E_i - E_m). \tag{8.104}$$

Note that the energy-shift $\mathit{\Delta}E_i$ is the same as that predicted by standard time-independent perturbation theory.

The probability of observing the system in state $|i\rangle$ at time $t > 0$, given that it is definately in state $|i\rangle$ at time $t = 0$, is given by

$$P_{i\to i}(t) = |c_i|^2 = \exp(-\mathit{\Gamma}_i\,t/\hbar), \tag{8.105}$$

where

$$\frac{\mathit{\Gamma}_i}{\hbar} = \sum_{m\neq i} w_{i\to m}. \tag{8.106}$$

Here, use has been made of Equation (8.79). Clearly, the rate of decay of the initial state is a simple function of the transition rates to the other states. Note that the system conserves probability up to second order in perturbed quantities, because

$$|c_i|^2 + \sum_{m\neq i}|c_m|^2 \simeq (1 - \mathit{\Gamma}_i\,t/\hbar) + \sum_{m\neq i} w_{i\to m}\,t = 1. \tag{8.107}$$

The quantity $\mathit{\Gamma}_i$ is called the *decay-width* of state $|i\rangle$. It is closely related to the mean lifetime of this state,

$$\tau_i = \frac{\hbar}{\mathit{\Gamma}_i}, \tag{8.108}$$

where

$$P_{i\to i} = \exp(-t/\tau_i). \tag{8.109}$$

According to Equation (8.101), the amplitude of state $|i\rangle$ both oscillates and decays as time progresses. Clearly, state $|i\rangle$ is not a stationary state in the presence of the time-dependent perturbation. However, we can still represent it as a superposition of stationary states (whose amplitudes simply oscillate in time). Thus,

$$\exp[-{\rm i}\,(E_i + \mathit{\Delta} E_i)\,t/\hbar]\,\exp(-\mathit{\Gamma}_i\,t/2\,\hbar) = \int dE\,f(E)\,\exp(-{\rm i}\,E\,t/\hbar), \tag{8.110}$$

where $f(E)$ is the weight of the stationary state with energy E in the superposition. The Fourier inversion theorem yields

$$|f(E)|^2 \propto \frac{1}{(E - [E_i + {\rm Re}(\mathit{\Delta}_i)])^2 + \mathit{\Gamma}_i^{\,2}/4}. \tag{8.111}$$

In the absence of the perturbation, $|f(E)|^2$ is basically a delta-function centered on the unperturbed energy E_i of state $|i\rangle$. In other words, state $|i\rangle$ is a stationary state whose energy is completely determined. In the presence of the perturbation, the energy of state $|i\rangle$ is *shifted* by ${\rm Re}(\mathit{\Delta}_i)$. The fact that the state is no longer stationary (i.e., it decays in time) implies that its energy cannot be exactly determined. Indeed, the energy of the state is smeared over some region of width (in energy) $\mathit{\Gamma}_i$ centered around the shifted energy $E_i+{\rm Re}(\mathit{\Delta}_i)$. The faster the decay of the state (i.e., the larger $\mathit{\Gamma}_i$), the more its energy is spread out. This effect is clearly a manifestation of the energy-time uncertainty relation $\mathit{\Delta} E\,\mathit{\Delta} t \sim \hbar$. One consequence of this effect is the existence of a *natural width* of spectral lines associated with the decay of some excited state to the ground state (or any other lower energy state). The uncertainty in energy of the excited state, due to its propensity to decay, gives rise to a slight smearing (in wavelength) of the spectral line associated with the transition. Strong lines, which correspond to fast transitions, are smeared out more that weak lines. For this reason, spectroscopists generally favor forbidden lines (see Section 8.10) for Doppler-shift measurements. Such lines are not as bright as those corresponding to allowed transitions, but they are a lot sharper.

8.8 Harmonic Perturbations

Consider a perturbation that oscillates sinusoidally in time. This is usually called a *harmonic perturbation*. Thus,

$$H_1(t) = V\,\exp({\rm i}\,\omega\,t) + V^\dagger\,\exp(-{\rm i}\,\omega\,t), \tag{8.112}$$

where V is, in general, a function of position, momentum, and spin operators.

Let us initiate the system in the eigenstate $|i\rangle$ of the unperturbed Hamiltonian, H_0, and switch on the harmonic perturbation at $t = 0$. It follows from Equation (8.58) that

$$\begin{aligned} c_n^{(1)} &= \frac{-{\rm i}}{\hbar}\int_0^t dt'\left[V_{ni}\,\exp({\rm i}\,\omega\,t') + V_{ni}^\dagger\,\exp(-{\rm i}\,\omega\,t')\right]\exp({\rm i}\,\omega_{ni}\,t') \\ &= \frac{1}{\hbar}\left(\frac{1-\exp[\,{\rm i}\,(\omega_{ni}+\omega)\,t]}{\omega_{ni}+\omega}\,V_{ni} + \frac{1-\exp[\,{\rm i}\,(\omega_{ni}-\omega)\,t]}{\omega_{ni}-\omega}\,V_{ni}^\dagger\right), \end{aligned} \tag{8.113}$$

where

$$V_{ni} = \langle n|\, V\, |i\rangle, \tag{8.114}$$

$$V^{\dagger}_{ni} = \langle n|\, V^{\dagger}\, |i\rangle = \langle i|\, V\, |n\rangle^{*}. \tag{8.115}$$

This formula is analogous to Equation (8.65), provided that

$$\omega_{ni} = \frac{E_n - E_i}{\hbar} \to \omega_{ni} \pm \omega. \tag{8.116}$$

Thus, it follows from the analysis of Section 8.6 that the transition probability $P_{i\to n}(t) = |c_n^{(1)}|^2$ is only appreciable in the limit $t \to \infty$ if

$$\omega_{ni} + \omega \simeq 0 \quad \text{or} \quad E_n \simeq E_i - \hbar\,\omega, \tag{8.117}$$

$$\omega_{ni} - \omega \simeq 0 \quad \text{or} \quad E_n \simeq E_i + \hbar\,\omega. \tag{8.118}$$

Clearly, (8.117) corresponds to the first term on the right-hand side of Equation (8.113), and (8.118) corresponds to the second term. The former term describes a process by which the system gives up energy $\hbar\,\omega$ to the perturbing field, while making a transition to a final state whose energy level is less than that of the initial state by $\hbar\,\omega$. This process is known as *stimulated emission*. The latter term describes a process by which the system gains energy $\hbar\,\omega$ from the perturbing field, while making a transition to a final state whose energy level exceeds that of the initial state by $\hbar\,\omega$. This process is known as *absorption*. In both cases, the total energy (i.e., that of the system *plus* the perturbing field) is conserved.

By analogy with Equation (8.78),

$$w_{i\to[n]} = \frac{2\pi}{\hbar}\,\overline{|V_{ni}|^2}\,\rho(E_n)\bigg|_{E_n=E_i-\hbar\,\omega}, \tag{8.119}$$

$$w_{i\to[n]} = \frac{2\pi}{\hbar}\,\overline{|V^{\dagger}_{ni}|^2}\,\rho(E_n)\bigg|_{E_n=E_i+\hbar\,\omega}. \tag{8.120}$$

Equation (8.119) specifies the transition rate for stimulated emission, whereas Equation (8.120) gives the transition rate for absorption. These equations are more usually written

$$w_{i\to n} = \frac{2\pi}{\hbar}\,|V_{ni}|^2\,\delta(E_n - E_i + \hbar\,\omega), \tag{8.121}$$

$$w_{i\to n} = \frac{2\pi}{\hbar}\,|V^{\dagger}_{ni}|^2\,\delta(E_n - E_i - \hbar\,\omega). \tag{8.122}$$

It is clear from Equations (8.114)–(8.115) that $|V^{\dagger}_{in}|^2 = |V_{ni}|^2$. It follows from Equations (8.119)–(8.120) that

$$\frac{w_{i\to[n]}}{\rho(E_n)} = \frac{w_{n\to[i]}}{\rho(E_i)}. \tag{8.123}$$

In other words, the rate of stimulated emission, divided by the density of final states for stimulated emission, equals the rate of absorption, divided by the density of final states for absorption. This result, which expresses a fundamental symmetry between absorption and stimulated emission, is known as *detailed balancing*, and is very important in statistical mechanics.

8.9 Absorption and Stimulated Emission of Radiation

Let us use some of the results of time-dependent perturbation theory to investigate the interaction of an atomic electron with classical (i.e., non-quantized) electromagnetic radiation.

The unperturbed Hamiltonian is

$$H_0 = \frac{p^2}{2\,m_e} + V_0(r). \tag{8.124}$$

The standard classical prescription for obtaining the Hamiltonian of a particle of charge q in the presence of an electromagnetic field is

$$\mathbf{p} \rightarrow \mathbf{p} - q\,\mathbf{A}, \tag{8.125}$$

$$H \rightarrow H - q\,\phi, \tag{8.126}$$

where $\mathbf{A}(\mathbf{r})$ is the vector potential and $\phi(\mathbf{r})$ is the scalar potential. Note that

$$\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t}, \tag{8.127}$$

$$\mathbf{B} = \nabla\times\mathbf{A}. \tag{8.128}$$

This prescription also works in quantum mechanics. Thus, the Hamiltonian of an atomic electron placed in an electromagnetic field is

$$H = \frac{|\mathbf{p} + e\,\mathbf{A}|^2}{2\,m_e} - e\,\phi + V_0(r), \tag{8.129}$$

where $\mathbf{A}$ and ϕ are real functions of the position operators. The above equation can be written

$$H = \frac{\left(p^2 + e\,\mathbf{A}\cdot\mathbf{p} + e\,\mathbf{p}\cdot\mathbf{A} + e^2A^2\right)}{2\,m_e} - e\,\phi + V_0(r). \tag{8.130}$$

Now,

$$\mathbf{p}\cdot\mathbf{A} = \mathbf{A}\cdot\mathbf{p}, \tag{8.131}$$

provided that we adopt the gauge $\nabla\cdot\mathbf{A} = 0$. Hence,

$$H = \frac{p^2}{2\,m_e} - \frac{e\,\mathbf{A}\cdot\mathbf{p}}{m_e} + \frac{e^2A^2}{2\,m_e} - e\,\phi + V_0(r). \tag{8.132}$$

Suppose that the perturbation corresponds to a monochromatic plane-wave, for which

$$\phi = 0, \tag{8.133}$$

$$\mathbf{A} = 2\,A_0\,\boldsymbol{\epsilon}\,\cos\left(\frac{\omega}{c}\,\mathbf{n}\cdot\mathbf{x} - \omega\,t\right), \tag{8.134}$$

where $\boldsymbol{\epsilon}$ and $\mathbf{n}$ are unit vectors that specify the direction of polarization and the direction of propagation, respectively. Note that $\boldsymbol{\epsilon}\cdot\mathbf{n} = 0$. The Hamiltonian becomes

$$H = H_0 + H_1(t), \tag{8.135}$$

with

$$H_0 = \frac{p^2}{2\,m_e} + V(r), \tag{8.136}$$

and

$$H_1 \simeq \frac{e\,\mathbf{A}\cdot\mathbf{p}}{m_e}, \tag{8.137}$$

where the A^2 term, which is second order in A_0, has been neglected.

The perturbing Hamiltonian can be written

$$H_1 = \frac{e\,A_0\,\boldsymbol{\epsilon}\cdot\mathbf{p}}{m_e}\left(\exp[\,\mathrm{i}\,(\omega/c)\,\mathbf{n}\cdot\mathbf{x} - \mathrm{i}\,\omega\,t] + \exp[-\mathrm{i}\,(\omega/c)\,\mathbf{n}\cdot\mathbf{x} + \mathrm{i}\,\omega\,t]\right). \tag{8.138}$$

This has the same form as Equation (8.112), provided that

$$V = -\frac{e\,A_0\,\boldsymbol{\epsilon}\cdot\mathbf{p}}{m_e}\,\exp[-\mathrm{i}\,(\omega/c)\,\mathbf{n}\cdot\mathbf{x}\,] \tag{8.139}$$

It is clear, by analogy with the previous analysis, that the first term on the right-hand side of Equation (8.138) describes the absorption of a photon of energy $\hbar\,\omega$, whereas the second term describes the stimulated emission of a photon of energy $\hbar\,\omega$. It follows from Equations (8.121) and (8.122) that the rates of absorption and stimulated emission are

$$w_{i\rightarrow n} = \frac{2\pi}{\hbar}\frac{e^2}{m_e^2}\,|A_0|^2\,|\langle n|\,\exp[\,\mathrm{i}\,(\omega/c)\,\mathbf{n}\cdot\mathbf{x}]\,\boldsymbol{\epsilon}\cdot\mathbf{p}\,|i\rangle|^2\,\delta(E_n - E_i - \hbar\,\omega), \tag{8.140}$$

and

$$w_{i\rightarrow n} = \frac{2\pi}{\hbar}\frac{e^2}{m_e^2}\,|A_0|^2\,|\langle n|\,\exp[-\mathrm{i}\,(\omega/c)\,\mathbf{n}\cdot\mathbf{x}]\,\boldsymbol{\epsilon}\cdot\mathbf{p}\,|i\rangle|^2\,\delta(E_n - E_i + \hbar\,\omega), \tag{8.141}$$

respectively.

Now, the energy density of a radiation field is

$$U = \frac{1}{2}\left(\frac{\epsilon_0\,E_0^2}{2} + \frac{B_0^2}{2\,\mu_0}\right), \tag{8.142}$$

where E_0 and $B_0 = E_0/c = 2\,A_0\,\omega/c$ are the peak electric and magnetic field-strengths, respectively. Hence,

$$U = 2\,\epsilon_0\,c\,\omega^2\,|A_0|^2, \tag{8.143}$$

and expressions (8.140) and (8.141) become

$$w_{i\rightarrow n} = \frac{\pi}{\hbar}\frac{e^2}{\epsilon_0\,m_e^2\,\omega^2}\,U\,|\langle n|\,\exp[\,\mathrm{i}\,(\omega/c)\,\mathbf{n}\cdot\mathbf{x}]\,\boldsymbol{\epsilon}\cdot\mathbf{p}\,|i\rangle|^2\,\delta(E_n - E_i - \hbar\,\omega), \tag{8.144}$$

and

$$w_{i\to n} = \frac{\pi}{\hbar}\,\frac{e^2}{\epsilon_0\, m_e^2\,\omega^2}\,U\,|\langle n|\,\exp[-{\rm i}\,(\omega/c)\,\mathbf{n}\cdot\mathbf{x}]\,\boldsymbol{\epsilon}\cdot\mathbf{p}\,|i\rangle|^2\,\delta(E_n - E_i + \hbar\,\omega), \tag{8.145}$$

respectively. Finally, if we imagine that the incident radiation has a range of different frequencies, so that

$$U = \int d\omega\, u(\omega), \tag{8.146}$$

where $d\omega\, u(\omega)$ is the energy density of radiation whose frequency lies in the range ω to $\omega + d\omega$, then we can integrate our transition rates over ω to give

$$w_{i\to n} = \frac{\pi}{\hbar^2}\,\frac{e^2}{\epsilon_0\, m_e^2\,\omega_{ni}^2}\,u(\omega_{ni})\,|\langle n|\,\exp[\,{\rm i}\,(\omega_{ni}/c)\,\mathbf{n}\cdot\mathbf{x}]\,\boldsymbol{\epsilon}\cdot\mathbf{p}\,|i\rangle|^2 \tag{8.147}$$

for absorption, and

$$w_{i\to n} = \frac{\pi}{\hbar^2}\,\frac{e^2}{\epsilon_0\, m_e^2\,\omega_{in}^2}\,u(\omega_{in})\,|\langle n|\,\exp[-{\rm i}\,(\omega_{in}/c)\,\mathbf{n}\cdot\mathbf{x}]\,\boldsymbol{\epsilon}\cdot\mathbf{p}\,|i\rangle|^2 \tag{8.148}$$

for stimulated emission. Here, $\omega_{ni} = (E_n - E_i)/\hbar > 0$ and $\omega_{in} = (E_i - E_n)/\hbar > 0$. Furthermore, we are assuming that the radiation is incoherent, so that intensities can be added.

8.10 Electric Dipole Approximation

In general, the wavelength of the type of electromagnetic radiation that induces, or is emitted during, transitions between different atomic energy levels is much larger than the typical size of a light atom. Thus,

$$\exp[\,{\rm i}\,(\omega/c)\,\mathbf{n}\cdot\mathbf{x}] = 1 + {\rm i}\,\frac{\omega}{c}\,\mathbf{n}\cdot\mathbf{x} + \cdots, \tag{8.149}$$

can be approximated by its first term, unity (remember that $\omega/c = 2\pi/\lambda$). This approximation is known as the *electric dipole approximation*. It follows that

$$\langle n|\,\exp[\,{\rm i}\,(\omega/c)\,\mathbf{n}\cdot\mathbf{x}]\,\boldsymbol{\epsilon}\cdot\mathbf{p}\,|i\rangle \simeq \boldsymbol{\epsilon}\cdot\langle n|\,\mathbf{p}\,|i\rangle. \tag{8.150}$$

It is readily demonstrated that

$$[\mathbf{x}, H_0] = \frac{{\rm i}\,\hbar\,\mathbf{p}}{m_e}, \tag{8.151}$$

so

$$\langle n|\,\mathbf{p}\,|i\rangle = -{\rm i}\,\frac{m_e}{\hbar}\langle n|\,[\mathbf{x}, H_0]\,|i\rangle = {\rm i}\,m_e\,\omega_{ni}\,\langle n|\,\mathbf{x}\,|i\rangle. \tag{8.152}$$

Thus, making use of the electric dipole approximation, we obtain

$$w_{i\to n} = 4\pi^2\,\alpha\,\frac{c}{\hbar}\,u(\omega_{ni})\,|\boldsymbol{\epsilon}\cdot\mathbf{f}_{ni}|^2 \tag{8.153}$$

for absorption, and

$$w_{i\to n} = 4\pi^2\,\alpha\,\frac{c}{\hbar}\,u(\omega_{in})\,|\boldsymbol{\epsilon}\cdot\mathbf{f}_{in}|^2 \tag{8.154}$$

for stimulated emission, where

$$\mathbf{f}_{ni} = \langle n|\,\mathbf{x}\,|i\rangle, \tag{8.155}$$

and $\alpha = e^2/(2\,\epsilon_0\,h\,c) \simeq 1/137$ is the fine structure constant.

Suppose that the radiation is polarized in the z-direction, so that $\boldsymbol{\epsilon} = \mathbf{e}_z$. We have already seen, from Section 7.4, that $\langle n|\,z\,|i\rangle = 0$ unless the initial and final states satisfy

$$\Delta l = \pm 1, \tag{8.156}$$

$$\Delta m = 0. \tag{8.157}$$

Here, l is the quantum number describing the total orbital angular momentum of the electron, and m is the quantum number describing the projection of the orbital angular momentum along the z-axis. It is easily demonstrated that $\langle n|\,x\,|i\rangle$ and $\langle n|\,y\,|i\rangle$ are only non-zero if

$$\Delta l = \pm 1, \tag{8.158}$$

$$\Delta m = \pm 1. \tag{8.159}$$

Thus, for generally directed radiation $\boldsymbol{\epsilon}\cdot\mathbf{f}_{ni}$ is only non-zero if

$$\Delta l = \pm 1, \tag{8.160}$$

$$\Delta m = 0, \pm 1. \tag{8.161}$$

These are termed the *selection rules* for electric dipole transitions. It is clear, for instance, that the electric dipole approximation allows a transition from a $2p$ state to a $1s$ state, but disallows a transition from a $2s$ to a $1s$ state. The latter transition is called a *forbidden transition*.

Forbidden transitions are not strictly forbidden. Instead, they take place at a far lower rate than transitions that are allowed according to the electric dipole approximation. After electric dipole transitions, the next most likely type of transition is a *magnetic dipole transition*, which is due to the interaction between the electron spin and the oscillating magnetic field of the incident electromagnetic radiation. Magnetic dipole transitions are typically about 10^5 times more unlikely than similar electric dipole transitions. The first-order term in Equation (8.149) yields so-called *electric quadrupole transitions*. These are typically about 10^8 times more unlikely than electric dipole transitions. Magnetic dipole and electric quadrupole transitions satisfy different selection rules than electric dipole transitions. For instance, the selection rules for electric quadrupole transitions are $\Delta l = 0, \pm 2$. Thus, transitions that are forbidden as electric dipole transitions may well be allowed as magnetic dipole or electric quadrupole transitions.

8.11 Spontaneous Emission

So far, we have calculated the rates of *radiation induced* transitions between two atomic states. This process is known as *absorption* when the energy of the final state exceeds that of the initial state, and *stimulated emission* when the energy of the final state is less than that of the initial state. Now, in the absence of any external radiation, we would not expect an atom in a given state to spontaneously jump into a state with a higher energy. On the other hand, it should be possible for such an atom to spontaneously jump into an state with a lower energy via the emission of a photon whose energy is equal to the difference between the energies of the initial and final states. This process is known as *spontaneous emission*.

It is possible to derive the rate of spontaneous emission between two atomic states from a knowledge of the corresponding absorption and stimulated emission rates using a famous thermodynamic argument due to Einstein. Consider a very large ensemble of similar atoms placed inside a closed cavity whose walls (which are assumed to be perfect emitters and absorbers of radiation) are held at the constant temperature T. Let the system have attained thermal equilibrium. According to statistical thermodynamics, the cavity is filled with so-called "black-body" electromagnetic radiation whose energy spectrum is

$$u(\omega) = \frac{\hbar}{\pi^2\,c^3}\,\frac{\omega^3}{\exp(\hbar\,\omega/k_B\,T)-1}, \tag{8.162}$$

where k_B is the Boltzmann constant. This well-known result was first obtained by Max Planck in 1900.

Consider two atomic states, labeled 2 and 1, with $E_2 > E_1$. One of the tenants of statistical thermodynamics is that in thermal equilibrium we have so-called *detailed balance*. This means that, irrespective of any other atomic states, the rate at which atoms in the ensemble leave state 2 due to transitions to state 1 is exactly balanced by the rate at which atoms enter state 2 due to transitions from state 1. The former rate (i.e., number of transitions per unit time in the ensemble) is written

$$W_{2\rightarrow 1} = N_2\,(w^{\rm spn}_{2\rightarrow 1} + w^{\rm stm}_{2\rightarrow 1}), \tag{8.163}$$

where $w^{\rm spn}_{2\rightarrow 1}$ and $w^{\rm arm}_{2\rightarrow 1}$ are the rates of spontaneous and stimulated emission, respectively, (for a single atom) between states 2 and 1, and N_2 is the number of atoms in the ensemble in state 2. Likewise, the latter rate takes the form

$$W_{1\rightarrow 2} = N_1\,w^{\rm abs}_{1\rightarrow 2}, \tag{8.164}$$

where $w^{\rm abs}_{1\rightarrow 2}$ is the rate of absorption (for a single atom) between states 1 and 2, and N_1 is the number of atoms in the ensemble in state 1. The above expressions describe how atoms in the ensemble make transitions from state 2 to state 1 due to a combination of spontaneous and stimulated emission, and make the opposite transition as a consequence of absorption. In thermal equilibrium, we have $W_{2\rightarrow 1} = W_{1\rightarrow 2}$, which gives

$$w^{\rm spn}_{2\rightarrow 1} = \frac{N_1}{N_2}\,w^{\rm abs}_{1\rightarrow 2} - w^{\rm stm}_{2\rightarrow 1}. \tag{8.165}$$

Equations (8.153) and (8.154) imply that

$$w^{\rm spn}_{i\rightarrow f} = \frac{4\pi^2\,\alpha\,c}{\hbar}\left(\frac{N_1}{N_2}-1\right)u(\omega_{21})\left\langle|\boldsymbol{\epsilon}\cdot{\bf f}_{21}|^2\right\rangle, \tag{8.166}$$

where $\omega_{21} = (E_2 - E_1)/\hbar$, and the large angle brackets denote an average over all possible directions of the incident radiation (because, in equilibrium, the radiation inside the cavity is isotropic). In fact, it is easily demonstrated that

$$\left\langle |\boldsymbol{\epsilon}\cdot\mathbf{f}_{21}|^2 \right\rangle = \frac{f_{21}^{\,2}}{3}, \tag{8.167}$$

where $f_{21}^{\,2}$ stands for

$$f_{21}^{\,2} = |\langle 2|\,x\,|1\rangle|^2 + |\langle 2|\,y\,|1\rangle|^2 + |\langle 2|\,z\,|1\rangle|^2. \tag{8.168}$$

Now, another famous result in statistical thermodynamics is that in thermal equilibrium the number of atoms in an ensemble occupying a state of energy E is proportional to $\exp(-E/k_B\,T)$. This implies that

$$\frac{N_1}{N_2} = \frac{\exp(-E_1/k_B\,T)}{\exp(-E_2/k_B\,T)} = \exp(\hbar\,\omega_{21}/k_B\,T). \tag{8.169}$$

Thus, it follows from Equations (8.162), (8.166), (8.167), and (8.169) that the rate of spontaneous emission between states 2 and 1 takes the form

$$w_{2\rightarrow 1}^{\rm spn} = \frac{\omega_{21}^{\,3}\,e^2\,f_{21}^{\,2}}{3\pi\,\epsilon_0\,\hbar\,c^3}. \tag{8.170}$$

Note, that, although the above result has been derived for an atom in a radiation-filled cavity, it remains correct even in the absence of radiation.

Let us estimate the typical value of the spontaneous emission rate for a hydrogen atom. We expect the matrix element f_{21} to be of order a_0, where a_0 is the Bohr radius. We also expect ω_{21} to be of order $|E_0|/\hbar$, where E_0 is the ground-state energy. It thus follows from Equation (8.170) that

$$w_{2\rightarrow 1}^{\rm spn} \sim \alpha^3\,\omega_{21}, \tag{8.171}$$

where $\alpha \simeq 1/137$ is the fine structure constant. This is an important result, because our perturbation expansion is based on the assumption that the transition rate between different energy eigenstates is much slower than the frequency of phase oscillation of these states: i.e., that $w_{2\rightarrow 1}^{\rm spn} \ll \omega_{21}$. This is indeed the case.

Exercises

8.1 Demonstrate that $\mathbf{p}\cdot\mathbf{A} = \mathbf{A}\cdot\mathbf{p}$ when $\nabla\cdot\mathbf{A} = 0$, where $\mathbf{p}$ is the momentum operator, and $\mathbf{A}(\mathbf{x})$ is a real function of the position operator, $\mathbf{x}$. Hence, show that the Hamiltonian (8.132) is Hermitian.

8.2 Find the selection rules for the matrix elements $\langle n,l,m|\,x\,|n',l',m'\rangle$, $\langle n,l,m|\,y\,|n',l',m'\rangle$, and $\langle n,l,m|\,z\,|n',l',m'\rangle$ to be non-zero. Here, $|n,l,m\rangle$ denotes an energy eigenket of a hydrogen-like atom corresponding to the conventional quantum numbers, n, l, and m.

8.3 Demonstrate that

$$\left\langle |\boldsymbol{\epsilon}\cdot\mathbf{f}_{21}|^2 \right\rangle = \frac{f_{21}^{\,2}}{3},$$

where the average is taken over all directions of the incident radiation.

8.4 Demonstrate that the spontaneous decay rate (via an electric dipole transition) from any 2p state to a 1s state of a hydrogen atom is

$$w_{2p\rightarrow 1s} = \left(\frac{2}{3}\right)^8 \alpha^5\,\frac{m_e\,c^2}{\hbar} = 6.26\times 10^8\,\mathrm{s}^{-1},$$

where α is the fine structure constant. Hence, deduce that the natural line width of the associated spectral line is

$$\frac{\Delta\lambda}{\lambda} \simeq 4\times 10^{-8}.$$

The only non-zero $1s \leftrightarrow 2p$ electric dipole matrix elements take the values

$$\langle 1,0,0|\,x\,|2,1,\pm 1\rangle = \pm\frac{2^7}{3^5}\,a_0,$$

$$\langle 1,0,0|\,y\,|2,1,\pm 1\rangle = \mathrm{i}\,\frac{2^7}{3^5}\,a_0,$$

$$\langle 1,0,0|\,z\,|2,1,0\rangle = \sqrt{2}\,\frac{2^7}{3^5}\,a_0,$$

where a_0 is the Bohr radius.

9 Scattering Theory

9.1 Introduction

Historically, data regarding quantum phenomena has been obtained from two main sources—the study of spectroscopic lines, and scattering experiments. We have already developed theories that account for some aspects of the spectra of hydrogen-like atoms. Let us now examine the quantum theory of scattering.

9.2 Fundamental Equations

Consider time-independent scattering theory, for which the Hamiltonian of the system is written

$$H = H_0 + H_1, \tag{9.1}$$

where H_0 is the Hamiltonian of a free particle of mass m,

$$H_0 = \frac{p^2}{2\,m}, \tag{9.2}$$

and H_1 represents the non-time-varying source of the scattering. Let $|\phi\rangle$ be an energy eigenket of H_0,

$$H_0\,|\phi\rangle = E\,|\phi\rangle, \tag{9.3}$$

whose wavefunction $\langle \mathbf{x}'|\phi\rangle$ is $\phi(\mathbf{x}')$. This state is assumed to be a plane wave state or, possibly, a spherical wave state. Schrödinger's equation for the scattering problem is

$$(H_0 + H_1)\,|\psi\rangle = E\,|\psi\rangle, \tag{9.4}$$

where $|\psi\rangle$ is an energy eigenstate of the total Hamiltonian whose wavefunction $\langle \mathbf{x}'|\psi\rangle$ is $\psi(\mathbf{x}')$. In general, both H_0 and $H_0 + H_1$ have continuous energy spectra: i.e., their energy eigenstates are unbound. We require a solution of Equation (9.4) that satisfies the boundary condition $|\psi\rangle \rightarrow |\phi\rangle$ as $H_1 \rightarrow 0$. Here, $|\phi\rangle$ is a solution of the free particle Schrödinger equation, (9.3), corresponding to the same energy eigenvalue.

Adopting the Schrödinger representation, we can write the scattering problem (9.4) in the form

$$(\nabla^2 + k^2)\,\psi(\mathbf{x}) = \frac{2\,m}{\hbar^2}\,\langle \mathbf{x}|\,H_1\,|\psi\rangle, \tag{9.5}$$

where

$$E = \frac{\hbar^2 k^2}{2\,m}. \tag{9.6}$$

Equation (9.5) is called the *Helmholtz equation*, and can be inverted using standard Green's function techniques. Thus,

$$\psi(\mathbf{x}) = \phi(\mathbf{x}) + \frac{2\,m}{\hbar^2}\int d^3x'\, G(\mathbf{x},\mathbf{x}')\,\langle\mathbf{x}'|\,H_1\,|\psi\rangle, \tag{9.7}$$

where

$$(\nabla^2 + k^2)\,G(\mathbf{x},\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}'). \tag{9.8}$$

Note that the solution (9.7) satisfies the boundary condition $|\psi\rangle \rightarrow |\phi\rangle$ as $H_1 \rightarrow 0$. As is well-known, the Green's function for the Helmholtz problem is given by

$$G(\mathbf{x},\mathbf{x}') = -\frac{\exp(\pm\mathrm{i}\,k\,|\mathbf{x}-\mathbf{x}'|\,)}{4\pi\,|\mathbf{x}-\mathbf{x}'|}. \tag{9.9}$$

Thus, Equation (9.7) becomes

$$\psi^{\pm}(\mathbf{x}) = \phi(\mathbf{x}) - \frac{2\,m}{\hbar^2}\int d^3x'\,\frac{\exp(\pm\mathrm{i}\,k\,|\mathbf{x}-\mathbf{x}'|\,)}{4\pi\,|\mathbf{x}-\mathbf{x}'|}\,\langle\mathbf{x}'|\,H_1\,|\psi^{\pm}\rangle. \tag{9.10}$$

Let us suppose that the scattering Hamiltonian, H_1, is only a function of the position operators. This implies that

$$\langle\mathbf{x}'|\,H_1\,|\mathbf{x}\rangle = V(\mathbf{x})\,\delta(\mathbf{x}-\mathbf{x}'). \tag{9.11}$$

We can write

$$\langle\mathbf{x}'|\,H_1\,|\psi^{\pm}\rangle = \int d^3x''\,\langle\mathbf{x}'|\,H_1\,|\mathbf{x}''\rangle\langle\mathbf{x}''|\psi^{\pm}\rangle\; = V(\mathbf{x}')\,\psi^{\pm}(\mathbf{x}'). \tag{9.12}$$

Thus, the integral equation (9.10) simplifies to

$$\psi^{\pm}(\mathbf{x}) = \phi(\mathbf{x}) - \frac{2\,m}{\hbar^2}\int d^3x'\,\frac{\exp(\pm\mathrm{i}\,k\,|\mathbf{x}-\mathbf{x}'|)}{4\pi\,|\mathbf{x}-\mathbf{x}'|}\,V(\mathbf{x}')\,\psi^{\pm}(\mathbf{x}'). \tag{9.13}$$

Suppose that the initial state $|\phi\rangle$ is a plane wave with wavevector $\mathbf{k}$ (i.e., a stream of particles of definite momentum $\mathbf{p} = \hbar\,\mathbf{k}$). The ket corresponding to this state is denoted $|\mathbf{k}\rangle$. The associated wavefunction takes the form

$$\langle\mathbf{x}|\mathbf{k}\rangle = \frac{\exp(\,\mathrm{i}\,\mathbf{k}\cdot\mathbf{x})}{(2\pi)^{3/2}}. \tag{9.14}$$

The wavefunction is normalized such that

$$\langle\mathbf{k}|\mathbf{k}'\rangle = \int d^3x\,\langle\mathbf{k}|\mathbf{x}\rangle\langle\mathbf{x}|\mathbf{k}'\rangle = \int d^3x\,\frac{\exp[-\mathrm{i}\,\mathbf{x}\cdot(\mathbf{k}-\mathbf{k}')]}{(2\pi)^3} = \delta(\mathbf{k}-\mathbf{k}'). \tag{9.15}$$

Suppose that the scattering potential $V(\mathbf{x})$ is only non-zero in some relatively localized region centered on the origin ($\mathbf{x} = \mathbf{0}$). Let us calculate the wavefunction $\psi(\mathbf{x})$ a long way from the scattering region. In other words, let us adopt the ordering $r \gg r'$. It is easily demonstrated that

$$|\mathbf{x}-\mathbf{x}'| \simeq r - \mathbf{e}_r\cdot\mathbf{x}' \tag{9.16}$$

to first order in r'/r, where

$$\mathbf{e}_r = \frac{\mathbf{x}}{r} \tag{9.17}$$

is a unit vector that points from the scattering region to the observation point. Here, $r = |\mathbf{x}|$ and $r' = |\mathbf{x}'|$. Let us define

$$\mathbf{k}' = k\,\mathbf{e}_r. \tag{9.18}$$

Clearly, $\mathbf{k}'$ is the wavevector for particles that possess the same energy as the incoming particles (i.e., $k' = k$), but propagate from the scattering region to the observation point. Note that

$$\exp(\pm\mathrm{i}\,k\,|\mathbf{x}-\mathbf{x}'|\,) \simeq \exp(\pm\mathrm{i}\,k\,r)\exp(\mp\mathrm{i}\,\mathbf{k}'\cdot\mathbf{x}'). \tag{9.19}$$

In the large-r limit, Equation (9.13) reduces to

$$\psi^{\pm}(\mathbf{x}) \simeq \frac{\exp(\,\mathrm{i}\,\mathbf{k}\cdot\mathbf{x})}{(2\pi)^{3/2}} - \frac{m}{2\pi\,\hbar^2}\frac{\exp(\pm\mathrm{i}\,k\,r)}{r}\int d^3x'\;\exp(\mp\mathrm{i}\,\mathbf{k}'\cdot\mathbf{x}')\,V(\mathbf{x}')\,\psi^{\pm}(\mathbf{x}'). \tag{9.20}$$

The first term on the right-hand side is the incident wave. The second term represents a spherical wave centred on the scattering region. The plus sign (on $\psi^{\pm}$) corresponds to a wave propagating away from the scattering region, whereas the minus sign corresponds to a wave propagating towards the scattering region. It is obvious that the former represents the physical solution. Thus, the wavefunction a long way from the scattering region can be written

$$\psi(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}}\left[\exp(\,\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}) + \frac{\exp(\,\mathrm{i}\,k\,r)}{r}f(\mathbf{k}',\mathbf{k})\right], \tag{9.21}$$

where

$$f(\mathbf{k}',\mathbf{k}) = -\frac{(2\pi)^2\,m}{\hbar^2}\int d^3\mathbf{x}'\;\frac{\exp(-\mathrm{i}\,\mathbf{k}'\cdot\mathbf{x}')}{(2\pi)^{3/2}}\,V(\mathbf{x}')\,\psi(\mathbf{x}') = -\frac{(2\pi)^2\,m}{\hbar^2}\,\langle\mathbf{k}'|\,H_1\,|\psi\rangle. \tag{9.22}$$

Let us define the differential cross-section, $d\sigma/d\Omega$, as the number of particles per unit time scattered into an element of solid angle $d\Omega$, divided by the incident flux of particles. Recall, from Chapter 3, that the probability current (i.e., the particle flux) associated with a wavefunction ψ is

$$\mathbf{j} = \frac{\hbar}{m}\,\mathrm{Im}(\psi^{*}\,\nabla\psi). \tag{9.23}$$

Thus, the probability flux associated with the incident wavefunction,

$$\frac{\exp(\,\mathrm{i}\,\mathbf{k}\cdot\mathbf{x})}{(2\pi)^{3/2}}, \tag{9.24}$$

is

$$\mathbf{j}_{\mathrm{inc}} = \frac{\hbar}{(2\pi)^3\,m}\,\mathbf{k}. \tag{9.25}$$

Likewise, the probability flux associated with the scattered wavefunction,

$$\frac{\exp(\,\mathrm{i}\,k\,r)}{(2\pi)^{3/2}}\frac{f(\mathbf{k}',\mathbf{k})}{r}, \tag{9.26}$$

is

$$\mathbf{j}_{\rm sca} = \frac{\hbar}{(2\pi)^3\,m}\frac{|f(\mathbf{k}',\mathbf{k})|^2}{r^2}\,k\,\mathbf{e}_r. \tag{9.27}$$

Now,

$$\frac{d\sigma}{d\Omega}\,d\Omega = \frac{r^2\,d\Omega\,|\mathbf{j}_{\rm sca}|}{|\mathbf{j}_{\rm inc}|}, \tag{9.28}$$

giving

$$\frac{d\sigma}{d\Omega} = |f(\mathbf{k}',\mathbf{k})|^2. \tag{9.29}$$

Thus, $|f(\mathbf{k}',\mathbf{k})|^2$ gives the differential cross-section for particles with incident momentum $\hbar\,\mathbf{k}$ to be scattered into states whose momentum vectors are directed in a range of solid angles $d\Omega$ about $\hbar\,\mathbf{k}'$. Note that the scattered particles possess the same energy as the incoming particles (i.e., $k' = k$). This is always the case for scattering Hamiltonians of the form specified in Equation (9.11).

9.3 Born Approximation

Equation (9.29) is not particularly useful, as it stands, because the quantity $f(\mathbf{k}',\mathbf{k})$ depends on the unknown ket $|\psi\rangle$. Recall that $\psi(\mathbf{x}) = \langle\mathbf{x}|\psi\rangle$ is the solution of the integral equation

$$\psi(\mathbf{x}) = \phi(\mathbf{x}) - \frac{m}{2\pi\,\hbar^2}\frac{\exp(\,\mathrm{i}\,k\,r)}{r}\int d^3x'\;\exp(-\mathrm{i}\,\mathbf{k}'\cdot\mathbf{x}')\,V(\mathbf{x}')\,\psi(\mathbf{x}'), \tag{9.30}$$

where $\phi(\mathbf{x})$ is the wavefunction of the incident state. According to the above equation, the total wavefunction is a superposition of the incident wavefunction and lots of spherical waves emitted from the scattering region. The strength of the spherical wave emitted at a given point is proportional to the local value of the scattering potential, V, as well as the local value of the wavefunction, ψ.

Suppose that the scattering is not particularly strong. In this case, it is reasonable to suppose that the total wavefunction, $\psi(\mathbf{x})$, does not differ substantially from the incident wavefunction, $\phi(\mathbf{x})$. Thus, we can obtain an expression for $f(\mathbf{k}',\mathbf{k})$ by making the substitution

$$\psi(\mathbf{x}) \rightarrow \phi(\mathbf{x}) = \frac{\exp(\,\mathrm{i}\,\mathbf{k}\cdot\mathbf{x})}{(2\pi)^{3/2}}. \tag{9.31}$$

This is called the *Born approximation.*

The Born approximation yields

$$f(\mathbf{k}',\mathbf{k}) \simeq -\frac{m}{2\pi\,\hbar^2}\int d^3x'\;\exp\left[\,\mathrm{i}\,(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}'\right]V(\mathbf{x}'). \tag{9.32}$$

Thus, $f(\mathbf{k}',\mathbf{k})$ is proportional to the Fourier transform of the scattering potential $V(\mathbf{x})$ with respect to the wavevector $\mathbf{q} \equiv \mathbf{k}-\mathbf{k}'$.

For a spherically symmetric potential,

$$f(\mathbf{k}',\mathbf{k}) \simeq -\frac{m}{2\pi\,\hbar^2}\int_0^\infty\int_0^\pi\int_0^{2\pi} dr'\,d\theta'\,d\phi'\;r'^{\,2}\,\sin\theta'\;\exp(\,\mathrm{i}\,q\,r'\cos\theta')\,V(r'), \tag{9.33}$$

giving

$$f(\mathbf{k}',\mathbf{k}) \simeq -\frac{2\,m}{\hbar^2\,q}\int_0^\infty dr'\,r'\,V(r')\sin(q\,r'). \tag{9.34}$$

Note that $f(\mathbf{k}',\mathbf{k})$ is just a function of q for a spherically symmetric potential. It is easily demonstrated that

$$q \equiv |\mathbf{k}-\mathbf{k}'| = 2\,k\,\sin(\theta/2), \tag{9.35}$$

where θ is the angle subtended between the vectors $\mathbf{k}$ and $\mathbf{k}'$. In other words, θ is the angle of scattering. Recall that the vectors $\mathbf{k}$ and $\mathbf{k}'$ have the same length, as a consequence of energy conservation.

Consider scattering by a Yukawa potential

$$V(r) = \frac{V_0\,\exp(-\mu\,r)}{\mu\,r}, \tag{9.36}$$

where V_0 is a constant, and $1/\mu$ measures the "range" of the potential. It follows from Equation (9.34) that

$$f(\theta) = -\frac{2\,m\,V_0}{\hbar^2\,\mu}\,\frac{1}{q^2+\mu^2}, \tag{9.37}$$

because

$$\int_0^\infty dr'\,\exp(-\mu\,r')\,\sin(q\,r') = \frac{q}{q^2+\mu^2}. \tag{9.38}$$

Thus, in the Born approximation, the differential cross-section for scattering by a Yukawa potential is

$$\frac{d\sigma}{d\Omega} \simeq \left(\frac{2\,m\,V_0}{\hbar^2\,\mu}\right)^2\frac{1}{[4\,k^2\,\sin^2(\theta/2)+\mu^2]^2}. \tag{9.39}$$

The Yukawa potential reduces to the familiar Coulomb potential as $\mu \to 0$, provided that $V_0/\mu \to Z\,Z'\,e^2/4\pi\,\epsilon_0$. In this limit, the Born differential cross-section becomes

$$\frac{d\sigma}{d\Omega} \simeq \left(\frac{2\,m\,Z\,Z'\,e^2}{4\pi\,\epsilon_0\,\hbar^2}\right)^2\frac{1}{16\,k^4\,\sin^4(\theta/2)}. \tag{9.40}$$

Recall that $\hbar\,k$ is equivalent to $|\mathbf{p}|$, so the above equation can be rewritten

$$\frac{d\sigma}{d\Omega} \simeq \left(\frac{Z\,Z'\,e^2}{16\pi\,\epsilon_0\,E}\right)^2\frac{1}{\sin^4(\theta/2)}, \tag{9.41}$$

where $E = p^2/2\,m$ is the kinetic energy of the incident particles. Equation (9.41) is identical to the classical Rutherford scattering cross-section formula.

The Born approximation is valid provided that $\psi(\mathbf{x})$ is not too different from $\phi(\mathbf{x})$ in the scattering region. It follows, from Equation (9.13), that the condition for $\psi(\mathbf{x}) \simeq \phi(\mathbf{x})$ in the vicinity of $\mathbf{x} = \mathbf{0}$ is

$$\left|\frac{m}{2\pi\,\hbar^2}\int d^3x'\,\frac{\exp(\,\mathrm{i}\,k\,r')}{r'}\,V(\mathbf{x}')\right| \ll 1. \tag{9.42}$$

Consider the special case of the Yukawa potential. At low energies, (i.e., $k \ll \mu$) we can replace $\exp(\mathrm{i}\,k\,r')$ by unity, giving

$$\frac{2\,m}{\hbar^2}\frac{|V_0|}{\mu^2} \ll 1 \tag{9.43}$$

as the condition for the validity of the Born approximation. The condition for the Yukawa potential to develop a bound state is

$$\frac{2\,m}{\hbar^2}\frac{|V_0|}{\mu^2} \geq 2.7, \tag{9.44}$$

where V_0 is negative. Thus, if the potential is strong enough to form a bound state then the Born approximation is likely to break down. In the high-k limit, Equation (9.42) yields

$$\frac{2\,m}{\hbar^2}\frac{|V_0|}{\mu\,k} \ll 1. \tag{9.45}$$

This inequality becomes progressively easier to satisfy as k increases, implying that the Born approximation is more accurate at high incident particle energies.

9.4 Partial Waves

We can assume, without loss of generality, that the incident wavefunction is characterized by a wavevector $\mathbf{k}$ that is aligned parallel to the z-axis. The scattered wavefunction is characterized by a wavevector $\mathbf{k}'$ that has the same magnitude as $\mathbf{k}$, but, in general, points in a different direction. The direction of $\mathbf{k}'$ is specified by the polar angle θ (i.e., the angle subtended between the two wavevectors), and an azimuthal angle φ about the z-axis. Equation (9.34) strongly suggests that for a spherically symmetric scattering potential [i.e., $V(\mathbf{x}) = V(r)$] the scattering amplitude is a function of θ only: i.e.,

$$f(\theta,\varphi) = f(\theta). \tag{9.46}$$

It follows that neither the incident wavefunction,

$$\phi(\mathbf{x}) = \frac{\exp(\mathrm{i}\,k\,z)}{(2\pi)^{3/2}} = \frac{\exp(\mathrm{i}\,k\,r\cos\theta)}{(2\pi)^{3/2}}, \tag{9.47}$$

nor the total wavefunction,

$$\psi(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}}\left[\exp(\mathrm{i}\,k\,r\cos\theta) + \frac{\exp(\mathrm{i}\,k\,r)\,f(\theta)}{r}\right], \tag{9.48}$$

depend on the azimuthal angle φ.

Outside the range of the scattering potential, both $\phi(\mathbf{x})$ and $\psi(\mathbf{x})$ satisfy the free space Schrödinger equation

$$(\nabla^2 + k^2)\,\psi = 0. \tag{9.49}$$

Consider the most general solution to this equation in spherical polar coordinates that does not depend on the azimuthal angle φ. Separation of variables yields

$$\psi(r,\theta) = \sum_{l=0,\infty} R_l(r)\,P_l(\cos\theta), \tag{9.50}$$

since the Legendre polynomials $P_l(\cos\theta)$ form a complete set in θ-space. The Legendre polynomials are related to the spherical harmonics introduced in Chapter 4 via

$$P_l(\cos\theta) = \sqrt{\frac{4\pi}{2\,l+1}}\,Y_{l\,0}(\theta,\varphi). \tag{9.51}$$

Equations (9.49) and (9.50) can be combined to give

$$r^2\frac{d^2R_l}{dr^2} + 2\,r\frac{dR_l}{dr} + [k^2\,r^2 - l\,(l+1)]\,R_l = 0. \tag{9.52}$$

The two independent solutions to this equation are the *spherical Bessel function*, $j_l(k\,r)$, and the *Neumann function*, $\eta_l(k\,r)$, where

$$j_l(y) = y^l\left(-\frac{1}{y}\frac{d}{dy}\right)^l\frac{\sin y}{y}, \tag{9.53}$$

$$\eta_l(y) = -y^l\left(-\frac{1}{y}\frac{d}{dy}\right)^l\frac{\cos y}{y}. \tag{9.54}$$

Note that spherical Bessel functions are well-behaved in the limit $y \to 0$, whereas Neumann functions become singular. The asymptotic behaviour of these functions in the limit $y \to \infty$ is

$$j_l(y) \to \frac{\sin(y - l\,\pi/2)}{y}, \tag{9.55}$$

$$\eta_l(y) \to -\frac{\cos(y - l\,\pi/2)}{y}. \tag{9.56}$$

We can write

$$\exp(\,\mathrm{i}\,k\,r\cos\theta) = \sum_{l=0,\infty} a_l\,j_l(k\,r)\,P_l(\cos\theta), \tag{9.57}$$

where the a_l are constants. Note there are no Neumann functions in this expansion, because they are not well-behaved as $r \to 0$. The Legendre polynomials are orthogonal,

$$\int_{-1}^{1} d\mu\,P_n(\mu)\,P_m(\mu) = \frac{\delta_{n\,m}}{n+1/2}, \tag{9.58}$$

so we can invert the above expansion to give

$$a_l\,j_l(k\,r) = (l+1/2)\int_{-1}^{1} d\mu\,\exp(\,\mathrm{i}\,k\,r\,\mu)\,P_l(\mu). \tag{9.59}$$

It is well-known that

$$j_l(y) = \frac{(-\mathrm{i})^l}{2}\int_{-1}^{1} d\mu\,\exp(\,\mathrm{i}\,y\,\mu)\,P_l(\mu), \tag{9.60}$$

where $l = 0, \infty$. Thus,

$$a_l = \mathrm{i}^l\,(2\,l+1), \tag{9.61}$$

giving

$$\exp(\mathrm{i}\,k\,r\cos\theta) = \sum_{l=0,\infty} \mathrm{i}^{l}\,(2\,l+1)\,j_l(k\,r)\,P_l(\cos\theta). \tag{9.62}$$

The above expression tells us how to decompose a plane wave into a series of spherical waves (or "partial waves").

The most general solution for the total wavefunction outside the scattering region is

$$\psi(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \sum_{l=0,\infty} \left[A_l\,j_l(k\,r) + B_l\,\eta_l(k\,r)\right] P_l(\cos\theta), \tag{9.63}$$

where the A_l and B_l are constants. Note that the Neumann functions are allowed to appear in this expansion, because its region of validity does not include the origin. In the large-r limit, the total wavefunction reduces to

$$\psi(\mathbf{x}) \simeq \frac{1}{(2\pi)^{3/2}} \sum_{l=0,\infty} \left[A_l\,\frac{\sin(k\,r - l\,\pi/2)}{k\,r} - B_l\,\frac{\cos(k\,r - l\,\pi/2)}{k\,r}\right] P_l(\cos\theta), \tag{9.64}$$

where use has been made of Equations (9.55)–(9.56). The above expression can also be written

$$\psi(\mathbf{x}) \simeq \frac{1}{(2\pi)^{3/2}} \sum_{l} C_l\,\frac{\sin(k\,r - l\,\pi/2 + \delta_l)}{k\,r}\,P_l(\cos\theta), \tag{9.65}$$

where the sine and cosine functions have been combined to give a sine function that is phase-shifted by δ_l.

Equation (9.65) yields

$$\psi(\mathbf{x}) \simeq \frac{1}{(2\pi)^{3/2}} \sum_{l} C_l\,\frac{\exp[\,\mathrm{i}\,(k\,r - l\,\pi/2 + \delta_l)] - \exp[-\mathrm{i}\,(k\,r - l\,\pi/2 + \delta_l)]}{2\,\mathrm{i}\,k\,r}\,P_l(\cos\theta), \tag{9.66}$$

which contains both incoming and outgoing spherical waves. What is the source of the incoming waves? Obviously, they must be part of the large-r asymptotic expansion of the incident wavefunction. In fact, it is easily seen that

$$\phi(\mathbf{x}) \simeq \frac{1}{(2\pi)^{3/2}} \sum_{l=0,\infty} \mathrm{i}^{l}\,(2l+1)\,\frac{\exp[\,\mathrm{i}\,(k\,r - l\,\pi/2)] - \exp[-\mathrm{i}\,(k\,r - l\,\pi/2)]}{2\,\mathrm{i}\,k\,r}\,P_l(\cos\theta), \tag{9.67}$$

in the large-r limit. Now, Equations (9.47) and (9.48) give

$$(2\pi)^{3/2}[\psi(\mathbf{x}) - \phi(\mathbf{x})] = \frac{\exp(\,\mathrm{i}\,k\,r)}{r}\,f(\theta). \tag{9.68}$$

Note that the right-hand side consists only of an outgoing spherical wave. This implies that the coefficients of the incoming spherical waves in the large-r expansions of $\psi(\mathbf{x})$ and $\phi(\mathbf{x})$ must be equal. It follows from Equations (9.66) and (9.67) that

$$C_l = (2\,l+1)\,\exp[\,\mathrm{i}\,(\delta_l + l\,\pi/2)]. \tag{9.69}$$

Thus, Equations (9.66)–(9.68) yield

$$f(\theta) = \sum_{l=0,\infty} (2\,l+1)\,\frac{\exp(\mathrm{i}\,\delta_l)}{k}\,\sin\delta_l\,P_l(\cos\theta). \tag{9.70}$$

Clearly, determining the scattering amplitude $f(\theta)$ via a decomposition into partial waves (i.e., spherical waves) is equivalent to determining the phase-shifts δ_l.

9.5 Optical Theorem

The differential scattering cross-section $d\sigma/d\Omega$ is simply the modulus squared of the scattering amplitude $f(\theta)$. The total cross-section is given by

$$\begin{aligned}\sigma_{\rm total} &= \oint d\Omega\,|f(\theta)|^2\\ &= \frac{1}{k^2}\oint d\varphi\int_{-1}^{1} d\mu \sum_l\sum_{l'}(2\,l+1)\,(2\,l'+1)\exp[\,\mathrm{i}\,(\delta_l-\delta_{l'})]\,\sin\delta_l\,\sin\delta_{l'}\,P_l(\mu)\,P_{l'}(\mu),\end{aligned} \tag{9.71}$$

where $\mu = \cos\theta$. It follows that

$$\sigma_{\rm total} = \frac{4\pi}{k^2}\sum_{l=0,\infty}(2\,l+1)\,\sin^2\delta_l, \tag{9.72}$$

where use has been made of Equation (9.58). A comparison of this result with Equation (9.70) yields

$$\sigma_{\rm total} = \frac{4\pi}{k}\,{\rm Im}\,[f(0)]\,, \tag{9.73}$$

since $P_l(1) = 1$. This result is known as the *optical theorem*. It is a reflection of the fact that the very existence of scattering requires scattering in the forward ($\theta = 0$) direction in order to interfere with the incident wave, and thereby reduce the probability current in this direction.

It is usual to write

$$\sigma_{\rm total} = \sum_{l=0,\infty}\sigma_l, \tag{9.74}$$

where

$$\sigma_l = \frac{4\pi}{k^2}\,(2\,l+1)\,\sin^2\delta_l \tag{9.75}$$

is the lth partial cross-section: i.e., the contribution to the total cross-section from the lth partial wave. Note that the maximum value for the lth partial cross-section occurs when the phase-shift δ_l takes the value $\pi/2$.

9.6 Determination of Phase-Shifts

Let us now consider how the phase-shifts δ_l can be evaluated. Consider a spherically symmetric potential $V(r)$ that vanishes for $r > a$, where a is termed the range of the potential. In the region $r > a$, the wavefunction $\psi(\mathbf{x})$ satisfies the free-space Schrödinger equation (9.49). The most general solution that is consistent with no incoming spherical waves is

$$\psi(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \sum_{l=0,\infty} \mathrm{i}^l\,(2\,l+1)\,A_l(r)\,P_l(\cos\theta), \tag{9.76}$$

where

$$A_l(r) = \exp(\,\mathrm{i}\,\delta_l)\,\left[\cos\delta_l\,j_l(k\,r) - \sin\delta_l\,\eta_l(k\,r)\right]. \tag{9.77}$$

Note that Neumann functions are allowed to appear in the above expression, because its region of validity does not include the origin (where $V \neq 0$). The logarithmic derivative of the lth radial wavefunction $A_l(r)$ just outside the range of the potential is given by

$$\beta_{l+} = k\,a\left[\frac{\cos\delta_l\,j_l'(k\,a) - \sin\delta_l\,\eta_l'(k\,a)}{\cos\delta_l\,j_l(k\,a) - \sin\delta_l\,\eta_l(k\,a)}\right], \tag{9.78}$$

where $j_l'(x)$ denotes $d\,j_l(x)/dx$, etc. The above equation can be inverted to give

$$\tan\delta_l = \frac{k\,a\,j_l'(k\,a) - \beta_{l+}\,j_l(k\,a)}{k\,a\,\eta_l'(k\,a) - \beta_{l+}\,\eta_l(k\,a)}. \tag{9.79}$$

Thus, the problem of determining the phase-shift δ_l is equivalent to that of determining β_{l+}.

The most general solution to Schrödinger's equation inside the range of the potential ($r < a$) that does not depend on the azimuthal angle φ is

$$\psi(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \sum_{l=0,\infty} \mathrm{i}^l\,(2\,l+1)\,R_l(r)\,P_l(\cos\theta), \tag{9.80}$$

where

$$R_l(r) = \frac{u_l(r)}{r}, \tag{9.81}$$

and

$$\frac{d^2u_l}{dr^2} + \left[k^2 - \frac{2m}{\hbar^2}\,V - \frac{l\,(l+1)}{r^2}\right]u_l = 0. \tag{9.82}$$

The boundary condition

$$u_l(0) = 0 \tag{9.83}$$

ensures that the radial wavefunction is well-behaved at the origin. We can launch a well-behaved solution of the above equation from $r = 0$, integrate out to $r = a$, and form the logarithmic derivative

$$\beta_{l-} = \frac{1}{(u_l/r)}\left.\frac{d(u_l/r)}{dr}\right|_{r=a}. \tag{9.84}$$

Because $\psi(\mathbf{x})$ and its first derivatives are necessarily continuous for physically acceptible wavefunctions, it follows that

$$\beta_{l+} = \beta_{l-}. \tag{9.85}$$

The phase-shift δ_l is obtainable from Equation (9.79).

9.7 Hard Sphere Scattering

Let us test out this scheme using a particularly simple example. Consider scattering by a hard sphere, for which the potential is infinite for $r < a$, and zero for $r > a$. It follows that $\psi(\mathbf{x})$ is zero in the region $r < a$, which implies that $u_l = 0$ for all l. Thus,

$$\beta_{l-} = \beta_{l+} = \infty, \tag{9.86}$$

for all l. It follows from Equation (9.79) that

$$\tan\delta_l = \frac{j_l(k\,a)}{\eta_l(k\,a)}. \tag{9.87}$$

Consider the $l = 0$ partial wave, which is usually referred to as the s-wave. Equation (9.87) yields

$$\tan\delta_0 = \frac{\sin(k\,a)/k\,a}{-\cos(k\,a)/ka} = -\tan(k\,a), \tag{9.88}$$

where use has been made of Equations (9.53)–(9.54). It follows that

$$\delta_0 = -k\,a. \tag{9.89}$$

The s-wave radial wave function is

$$A_0(r) = \exp(-\mathrm{i}\,k\,a)\left[\frac{\cos(k\,a)\,\sin(k\,r) - \sin(k\,a)\,\cos(k\,r)}{k\,r}\right] = \exp(-\mathrm{i}\,k\,a)\,\frac{\sin[k\,(r-a)]}{k\,r}. \tag{9.90}$$

The corresponding radial wavefunction for the incident wave takes the form

$$\tilde{A}_0(r) = \frac{\sin(k\,r)}{k\,r}. \tag{9.91}$$

It is clear that the actual $l = 0$ radial wavefunction is similar to the incident $l = 0$ wavefunction, except that it is phase-shifted by $k\,a$.

Let us consider the low and high energy asymptotic limits of $\tan\delta_l$. Low energy corresponds to $k\,a \ll 1$. In this limit, the spherical Bessel functions and Neumann functions reduce to:

$$j_l(k\,r) \simeq \frac{(k\,r)^l}{(2\,l+1)!!}, \tag{9.92}$$

$$\eta_l(k\,r) \simeq -\frac{(2\,l-1)!!}{(k\,r)^{l+1}}, \tag{9.93}$$

where $n!! = n\,(n-2)\,(n-4)\cdots 1$. It follows that

$$\tan\delta_l = \frac{-(k\,a)^{2\,l+1}}{(2\,l+1)\,[(2\,l-1)!!]^2}. \tag{9.94}$$

It is clear that we can neglect δ_l, with $l > 0$, with respect to δ_0. In other words, at low energy only s-wave scattering (i.e., spherically symmetric scattering) is important. It follows from Equations (9.29), (9.70), and (9.89) that

$$\frac{d\sigma}{d\Omega} = \frac{\sin^2(k\,a)}{k^2} \simeq a^2 \tag{9.95}$$

for $k\,a \ll 1$. Note that the total cross-section

$$\sigma_{\rm total} = \oint d\Omega\,\frac{d\sigma}{d\Omega} = 4\pi\,a^2 \tag{9.96}$$

is *four times* the *geometric cross-section* $\pi\,a^2$ (i.e., the cross-section for classical particles bouncing off a hard sphere of radius a). However, low energy scattering implies relatively long wavelengths, so we do not expect to obtain the classical result in this limit.

Consider the high energy limit $k\,a \gg 1$. At high energies, all partial waves up to $l_{\rm max} = k\,a$ contribute significantly to the scattering cross-section. It follows from Equation (9.72) that

$$\sigma_{\rm total} = \frac{4\pi}{k^2}\sum_{l=0,l_{\rm max}}(2\,l+1)\,\sin^2\delta_l. \tag{9.97}$$

With so many l values contributing, it is legitimate to replace $\sin^2\delta_l$ by its average value $1/2$. Thus,

$$\sigma_{\rm total} = \sum_{l=0,k\,a}\frac{2\pi}{k^2}\,(2\,l+1) \simeq 2\pi\,a^2. \tag{9.98}$$

This is *twice* the classical result, which is somewhat surprizing, because we might expect to obtain the classical result in the short wavelength limit. For hard sphere scattering, incident waves with impact parameters less than a must be deflected. However, in order to produce a "shadow" behind the sphere there must be scattering in the forward direction (recall the optical theorem) to produce destructive interference with the incident plane wave. In fact, the interference is not completely destructive, and the shadow has a bright spot in the forward direction. The effective cross-section associated with this bright spot is $\pi\,a^2$ which, when combined with the cross-section for classical reflection, $\pi\,a^2$, gives the actual cross-section of $2\pi\,a^2$.

9.8 Low Energy Scattering

At low energies (i.e., when $1/k$ is much larger than the range of the potential) partial waves with $l > 0$, in general, make a negligible contribution to the scattering cross-section. It follows that, at these energies, with a finite range potential, only s-wave scattering is important.

As a specific example, let us consider scattering by a finite potential well, characterized by $V = V_0$ for $r < a$, and $V = 0$ for $r \geq a$. Here, V_0 is a constant. The potential is repulsive for $V_0 > 0$, and attractive for $V_0 < 0$. The external wavefunction is given by [see Equation (9.77)]

$$A_0(r) = \exp(\,{\rm i}\,\delta_0)\,\left[j_0(k\,r)\cos\delta_0 - \eta_0(k\,r)\sin\delta_0\right] = \frac{\exp(\,{\rm i}\,\delta_0)\,\sin(k\,r+\delta_0)}{k\,r}, \tag{9.99}$$

where use has been made of Equations (9.53)–(9.54). The internal wavefunction follows from Equation (9.82). We obtain

$$A_0(r) = B\,\frac{\sin(k'\,r)}{r}, \tag{9.100}$$

where use has been made of the boundary condition (9.83). Here, B is a constant, and

$$E - V_0 = \frac{\hbar^2\,k'^{\,2}}{2\,m}. \tag{9.101}$$

Note that Equation (9.100) only applies when $E > V_0$. For $E < V_0$, we have

$$A_0(r) = B\,\frac{\sinh(\kappa\,r)}{r}, \tag{9.102}$$

where

$$V_0 - E = \frac{\hbar^2\kappa^2}{2\,m}. \tag{9.103}$$

Matching $A_0(r)$, and its radial derivative at $r = a$, yields

$$\tan(k\,a + \delta_0) = \frac{k}{k'}\,\tan(k'\,a) \tag{9.104}$$

for $E > V_0$, and

$$\tan(k\,a + \delta_0) = \frac{k}{\kappa}\,\tanh(\kappa\,a) \tag{9.105}$$

for $E < V_0$.

Consider an attractive potential, for which $E > V_0$. Suppose that $|V_0| \gg E$ (i.e., the depth of the potential well is much larger than the energy of the incident particles), so that $k' \gg k$. It follows from Equation (9.104) that, unless $\tan(k'\,a)$ becomes extremely large, the right-hand side is much less that unity, so replacing the tangent of a small quantity with the quantity itself, we obtain

$$k\,a + \delta_0 \simeq \frac{k}{k'}\,\tan(k'\,a). \tag{9.106}$$

This yields

$$\delta_0 \simeq k\,a\left[\frac{\tan(k'\,a)}{k'\,a} - 1\right]. \tag{9.107}$$

According to Equation (9.97), the scattering cross-section is given by

$$\sigma_{\rm total} \simeq \frac{4\pi}{k^2}\,\sin^2\delta_0 = 4\pi\,a^2\left[\frac{\tan(k'\,a)}{(k'\,a)} - 1\right]^2. \tag{9.108}$$

Now,

$$k'a = \sqrt{k^2\,a^2 + \frac{2\,m\,|V_0|\,a^2}{\hbar^2}}, \tag{9.109}$$

so for sufficiently small values of $k\,a$,

$$k'a \simeq \sqrt{\frac{2\,m\,|V_0|\,a^2}{\hbar^2}}. \tag{9.110}$$

It follows that the total (s-wave) scattering cross-section is independent of the energy of the incident particles (provided that this energy is sufficiently small).

Note that there are values of $k'a$ (e.g., $k'a \simeq 4.49$) at which $\delta_0 \to \pi$, and the scattering cross-section (9.108) vanishes, despite the very strong attraction of the potential. In reality, the cross-section is not exactly zero, because of contributions from $l > 0$ partial waves. But, at low incident energies, these contributions are small. It follows that there are certain values of V_0 and k that give rise to almost perfect transmission of the incident wave. This is called the *Ramsauer-Townsend effect*, and has been observed experimentally.

9.9 Resonance Scattering

There is a significant exception to the independence of the cross-section on energy. Suppose that the quantity $\sqrt{2\,m\,|V_0|\,a^2/\hbar^2}$ is slightly less than $\pi/2$. As the incident energy increases, $k'a$, which is given by Equation (9.109), can reach the value $\pi/2$. In this case, $\tan(k'\,a)$ becomes infinite, so we can no longer assume that the right-hand side of Equation (9.104) is small. In fact, at the value of the incident energy when $k'a = \pi/2$, it follows from Equation (9.104) that $k\,a + \delta_0 = \pi/2$, or $\delta_0 \simeq \pi/2$ (because we are assuming that $k\,a \ll 1$). This implies that

$$\sigma_{\rm total} = \frac{4\pi}{k^2}\sin^2\delta_0 = 4\pi\,a^2\left(\frac{1}{k^2\,a^2}\right). \tag{9.111}$$

Note that the cross-section now depends on the energy. Furthermore, the magnitude of the cross-section is much larger than that given in Equation (9.108) for $k'a \neq \pi/2$ (since $k\,a \ll 1$).

The origin of this rather strange behaviour is quite simple. The condition

$$\sqrt{\frac{2\,m\,|V_0|\,a^2}{\hbar^2}} = \frac{\pi}{2} \tag{9.112}$$

is equivalent to the condition that a spherical well of depth V_0 possesses a bound state at zero energy. Thus, for a potential well that satisfies the above equation, the energy of the scattering system is essentially the same as the energy of the bound state. In this situation, an incident particle would like to form a bound state in the potential well. However, the bound state is not stable, because the system has a small positive energy. Nevertheless, this sort of *resonance scattering* is best understood as the capture of an incident particle to form a metastable bound state, and the subsequent decay of the bound state and release of the particle. The cross-section for resonance scattering is generally far higher than that for non-resonance scattering.

We have seen that there is a resonant effect when the phase-shift of the s-wave takes the value $\pi/2$. There is nothing special about the $l = 0$ partial wave, so it is reasonable to assume that there is a similar resonance when the phase-shift of the lth partial wave is $\pi/2$. Suppose that δ_l attains the value $\pi/2$ at the incident energy E_0, so that

$$\delta_l(E_0) = \frac{\pi}{2}. \tag{9.113}$$

Let us expand $\cot\delta_l$ in the vicinity of the resonant energy:

$$\cot\delta_l(E) = \cot\delta_l(E_0) + \left(\frac{d\cot\delta_l}{dE}\right)_{E=E_0}(E-E_0) + \cdots = -\left(\frac{1}{\sin^2\delta_l}\frac{d\delta_l}{dE}\right)_{E=E_0}(E-E_0) + \cdots\,. \tag{9.114}$$

Defining

$$\left(\frac{d\delta_l(E)}{dE}\right)_{E=E_0} = \frac{2}{\Gamma}, \tag{9.115}$$

we obtain

$$\cot\delta_l(E) = -\frac{2}{\Gamma}(E-E_0) + \cdots\,. \tag{9.116}$$

Recall, from Equation (9.75), that the contribution of the lth partial wave to the scattering cross-section is

$$\sigma_l = \frac{4\pi}{k^2}(2\,l+1)\,\sin^2\delta_l = \frac{4\pi}{k^2}(2\,l+1)\,\frac{1}{1+\cot^2\delta_l}. \tag{9.117}$$

Thus,

$$\sigma_l \simeq \frac{4\pi}{k^2}(2\,l+1)\,\frac{\Gamma^2/4}{(E-E_0)^2+\Gamma^2/4}. \tag{9.118}$$

This is the famous *Breit-Wigner* formula. The variation of the partial cross-section σ_l with the incident energy has the form of a classical resonance curve. The quantity Γ is the width of the resonance (in energy). We can interpret the Breit-Wigner formula as describing the absorption of an incident particle to form a metastable state, of energy E_0, and lifetime $\tau = \hbar/\Gamma$.

Exercises

9.1 Consider a scattering potential of the form

$$V(r) = V_0\,\exp(-r^2/a^2).$$

Calculate the differential scattering cross-section, $d\sigma/d\Omega$, using the Born approximation.

9.2 Consider a scattering potential that takes the constant value V_0 for $r < R$, and is zero for $r > R$, where V_0 may be either positive or negative. Using the method of partial waves, show that for $|V_0| \ll E = \hbar^2\,k^2/2\,m$, and $k\,R \ll 1$, the differential cross-section is isotropic, and that the total cross-section is

$$\sigma_{\rm tot} = \left(\frac{16\pi}{9}\right)\frac{m^2\,V_0^{\,2}\,R^6}{\hbar^4}.$$

Suppose that the energy is slightly raised. Show that the angular distribution can then be written in the form

$$\frac{d\sigma}{d\Omega} = A + B\,\cos\theta.$$

Obtain an approximate expression for B/A.

9.3 Consider scattering by a repulsive δ-shell potential:

$$V(r) = \left(\frac{\hbar^2}{2\,m}\right)\gamma\,\delta(r-R),$$

where $\gamma > 0$. Find the equation that determines the s-wave phase-shift, δ_0, as a function of k (where $E = \hbar^2\,k^2/2\,m$). Assume that $\gamma \gg R^{-1}$, k. Show that if $\tan(k\,R)$ is not close to zero then the s-wave phase-shift resembles the hard sphere result discussed in the text. Furthermore, show that if $\tan(k\,R)$ is close to zero then resonance behavior is possible: i.e., $\cot\delta_0$ goes through zero from the positive side as k increases. Determine the approximate positions of the resonances (retaining terms up to order $1/\gamma$). Compare the resonant energies with the bound state energies for a particle confined within an infinite spherical well of radius R. Obtain an approximate expression for the resonance width

$$\Gamma = -\frac{2}{[d(\cot\delta_0)/dE]_{E=E_r}}.$$

Show that the resonances become extremely sharp as $\gamma \rightarrow \infty$.

9.4 Show that the differential cross-section for the elastic scattering of a fast electron by the ground-state of a hydrogen atom is

$$\frac{d\sigma}{d\Omega} = \left(\frac{2\,m_e\,e^2}{4\pi\,\epsilon_0\,\hbar^2\,q^2}\right)^2\left(1 - \frac{16}{[4+(q\,a_0)^2]^2}\right),$$

where $q = |\mathbf{k}-\mathbf{k}'|$, and a_0 is the Bohr radius.

10 Identical Particles

10.1 Permutation Symmetry

Consider a system consisting of a collection of identical particles. In classical mechanics, it is, in principle, possible to continuously monitor the position of each particle as a function of time. Hence, the constituent particles can be unambiguously labeled. In quantum mechanics, on the other hand, this is not possible because continuous position measurements would disturb the system. It follows that identical particles cannot be unambiguously labeled in quantum mechanics.

Consider a quantum system consisting of two identical particles. Suppose that one of the particles—particle 1, say—is characterized by the state ket $|k'\rangle$. Here, k' represents the eigenvalues of the complete set of commuting observables associated with the particle. Suppose that the other particle—particle 2—is characterized by the state ket $|k''\rangle$. The state ket for the whole system can be written in the product form

$$|k'\rangle\,|k''\rangle, \tag{10.1}$$

where it is understood that the first ket corresponds to particle 1, and the second to particle 2. We can also consider the ket

$$|k''\rangle\,|k'\rangle, \tag{10.2}$$

which corresponds to a state in which particle 1 has the eigenvalues k'', and particle 2 the eigenvalues k'.

Suppose that we were to measure all of the simultaneously measurable properties of our two-particle system. We might obtain the results k' for one particle, and k'' for the other. However, we have no way of knowing whether the corresponding state ket is $|k'\rangle\,|k''\rangle$ or $|k''\rangle\,|k'\rangle$, or any linear combination of these two kets. In other words, all state kets of the form

$$c_1\,|k'\rangle\,|k''\rangle + c_2\,|k''\rangle\,|k'\rangle, \tag{10.3}$$

correspond to an identical set of results when the properties of the system are measured. This phenomenon is known as *exchange degeneracy*. Such degeneracy is problematic because the specification of a complete set of observable eigenvalues in a system of identical particles does not seem to uniquely determine the corresponding state ket. Fortunately, nature has a way of avoiding this difficulty.

Consider the permutation operator P_{12}, which is defined such that

$$P_{12}\,|k'\rangle\,|k''\rangle = |k''\rangle\,|k'\rangle. \tag{10.4}$$

In other words, P_{12} swaps the identities of particles 1 and 2. It is easily seen that

$$P_{21} = P_{12}, \tag{10.5}$$

$$P_{12}^2 = 1. \tag{10.6}$$

Now, the Hamiltonian of a system of two identical particles must necessarily be a symmetric function of each particle's observables (because exchange of identical particles could not possibly affect the overall energy of the system). For instance,

$$H = \frac{\mathbf{p}_1^2}{2\,m} + \frac{\mathbf{p}_2^2}{2\,m} + V_{\rm pair}(|\mathbf{x}_1 - \mathbf{x}_2|) + V_{\rm ext}(\mathbf{x}_1) + V_{\rm ext}(\mathbf{x}_2). \tag{10.7}$$

Here, we have separated the mutual interaction of the two particles from their interaction with an external potential. It follows that if

$$H\,|k'\rangle\,|k''\rangle = E\,|k'\rangle\,|k''\rangle \tag{10.8}$$

then

$$H\,|k''\rangle\,|k'\rangle = E\,|k''\rangle\,|k'\rangle, \tag{10.9}$$

where E is the total energy. Operating on both sides of (10.8) with P_{12}, and making use of Equation (10.6), we obtain

$$P_{12}\,H\,P_{12}^2\,|k'\rangle\,|k''\rangle = E\,P_{12}\,|k'\rangle\,|k''\rangle, \tag{10.10}$$

or

$$P_{12}\,H\,P_{12}\,|k''\rangle\,|k'\rangle = E\,|k''\rangle\,|k'\rangle = H\,|k''\rangle\,|k'\rangle, \tag{10.11}$$

where use has been made of (10.9). We deduce that

$$P_{12}\,H\,P_{12} = H, \tag{10.12}$$

which implies [from (10.6)] that

$$[H, P_{12}] = 0. \tag{10.13}$$

In other words, an eigenstate of the Hamiltonian is a simultaneous eigenstate of the permutation operator P_{12}.

Now, according to Equation (10.6), the permutation operator possesses the eigenvalues +1 and −1, respectively. The corresponding properly normalized eigenstates are

$$|k'\,k''\rangle_+ = \frac{1}{\sqrt{2}}\left(|k'\rangle\,|k''\rangle + |k''\rangle\,|k'\rangle\right), \tag{10.14}$$

and

$$|k'\,k''\rangle_- = \frac{1}{\sqrt{2}}\left(|k'\rangle\,|k''\rangle - |k''\rangle\,|k'\rangle\right). \tag{10.15}$$

Here, it is assumed that $\langle k'|k''\rangle = \delta_{k'\,k''}$. Note that $|k'\,k''\rangle_+$ is symmetric with respect to interchange of particles—i.e.,

$$|k''\,k'\rangle_+ = +|k'\,k''\rangle_+, \tag{10.16}$$

whereas $|k'\,k''\rangle_-$ is antisymmetric—i.e.,

$$|k''\,k'\rangle_- = -|k'\,k''\rangle_-. \tag{10.17}$$

Let us now consider a system of three identical particles. We can represent the overall state ket as

$$|k'\,k''\,k'''\rangle, \tag{10.18}$$

where k', k'', and k''' are the eigenvalues of particles 1, 2, and 3, respectively. We can also define the two-particle permutation operators

$$P_{12}\,|k'\,k''\,k'''\rangle = |k''\,k'\,k'''\rangle, \tag{10.19}$$

$$P_{23}\,|k'\,k''\,k'''\rangle = |k'\,k'''\,k''\rangle, \tag{10.20}$$

$$P_{31}\,|k'\,k''\,k'''\rangle = |k'''\,k''\,k'\rangle. \tag{10.21}$$

It is easily demonstrated that

$$P_{21} = P_{12}, \tag{10.22}$$

$$P_{32} = P_{23}, \tag{10.23}$$

$$P_{13} = P_{31}, \tag{10.24}$$

and

$$P_{12}^2 = 1, \tag{10.25}$$

$$P_{23}^2 = 1, \tag{10.26}$$

$$P_{31}^2 = 1. \tag{10.27}$$

As before, the Hamiltonian of the system must be a symmetric function of the particle's observables: i.e.,

$$H\,|k'\,k''\,k'''\rangle = E\,|k'\,k''\,k'''\rangle, \tag{10.28}$$

$$H\,|k''\,k'''\,k'\rangle = E\,|k''\,k'''\,k'\rangle, \tag{10.29}$$

$$H\,|k'''\,k'\,k''\rangle = E\,|k'''\,k'\,k''\rangle, \tag{10.30}$$

$$H\,|k''\,k'\,k'''\rangle = E\,|k''\,k'\,k'''\rangle, \tag{10.31}$$

$$H\,|k'\,k'''\,k''\rangle = E\,|k'\,k'''\,k''\rangle, \tag{10.32}$$

$$H\,|k'''\,k''\,k'\rangle = E\,|k'''\,k''\,k'\rangle, \tag{10.33}$$

where E is the total energy. Using analogous arguments to those employed for the two-particle system, we deduce that

$$[H, P_{12}] = [H, P_{23}] = [H, P_{31}] = 0. \tag{10.34}$$

Hence, an eigenstate of the Hamiltonian is a simultaneous eigenstate of the permutation operators P_{12}, P_{23}, and P_{31}. However, according to Equations (10.25)–(10.27), the possible eigenvalues of these operators are ± 1.

Let us define the cyclic permutation operator P_{123}, where

$$P_{123}\,|k'\,k''\,k'''\rangle = |k'''\,k'\,k''\rangle. \tag{10.35}$$

It follows that

$$P_{123} = P_{12}\,P_{31} = P_{23}\,P_{12} = P_{31}\,P_{23}. \tag{10.36}$$

It is also clear from Equations (10.28) and (10.33) that

$$[H, P_{123}] = 0. \tag{10.37}$$

Thus, an eigenstate of the Hamiltonian is a simultaneous eigenstate of the permutation operators P_{12}, P_{23}, P_{31}, and P_{123}. Let λ_{12}, λ_{23}, λ_{31} and λ_{123} represent the eigenvalues of these operators, respectively. We know that $\lambda_{12} = \pm 1$, $\lambda_{23} = \pm 1$, and $\lambda_{31} = \pm 1$. Moreover, it follows from (10.36) that

$$\lambda_{123} = \lambda_{12}\,\lambda_{31} = \lambda_{23}\,\lambda_{12} = \lambda_{31}\,\lambda_{23}. \tag{10.38}$$

The above equations imply that

$$\lambda_{123} = +1, \tag{10.39}$$

and either

$$\lambda_{12} = \lambda_{23} = \lambda_{31} = +1, \tag{10.40}$$

or

$$\lambda_{12} = \lambda_{23} = \lambda_{31} = -1. \tag{10.41}$$

In other words, the multi-particle state ket must be either totally symmetric, or totally antisymmetric, with respect to swapping the identities of any given pair of particles. Thus, in terms of properly normalized single particle kets, the properly normalized totally symmetric and totally antisymmetric kets are

$$\begin{aligned}|k'\,k''\,k'''\rangle_+ = \frac{1}{\sqrt{3!}}\,\big(&|k'\rangle\,|k''\rangle\,|k'''\rangle + |k'''\rangle\,|k'\rangle\,|k''\rangle + |k''\rangle\,|k'''\rangle\,|k'\rangle \\ &+|k'''\rangle\,|k''\rangle\,|k'\rangle + |k'\rangle\,|k'''\rangle\,|k''\rangle + |k''\rangle\,|k'\rangle\,|k'''\rangle\big)\,,\end{aligned} \tag{10.42}$$

and

$$\begin{aligned}|k'\,k''\,k'''\rangle_- = \frac{1}{\sqrt{3!}}\,\big(&|k'\rangle\,|k''\rangle\,|k'''\rangle + |k'''\rangle\,|k'\rangle\,|k''\rangle + |k''\rangle\,|k'''\rangle\,|k'\rangle \\ &-|k'''\rangle\,|k''\rangle\,|k'\rangle - |k'\rangle\,|k'''\rangle\,|k''\rangle - |k''\rangle\,|k'\rangle\,|k'''\rangle\big)\,,\end{aligned} \tag{10.43}$$

respectively.

The above arguments can be generalized to systems of more than three identical particles in a straightforward manner.

10.2 Symmetrization Postulate

We have seen that the exchange degeneracy of a system of identical particles is such that a specification of a complete set of observable eigenvalues does not uniquely determine the corresponding state ket. However, we have also seen that there are only two possible state kets: i.e., a ket that is totally symmetric with respect to particle interchange, or a ket that is totally antisymmetric. It turns out that systems of identical particles possessing integer-spin (e.g., spin 0, or spin 1) always choose the totally symmetric ket, whereas systems of identical particles possessing half-integer-spin (e.g., spin 1/2) always choose the totally antisymmetric ket. This additional piece of information ensures that the specification of a complete set of observable eigenvalues of a system of identical particles does, in fact, uniquely determine the corresponding state ket.

Systems of identical particles whose state kets are totally symmetric with respect to particle interchange are said to obey *Bose-Einstein statistics*. Moreover, such particles are termed *bosons*. On the other hand, systems of identical particles whose state kets are totally antisymmetric with respect to particle interchange are said to obey *Fermi-Dirac statistics*, and the constituent particles are called *fermions*. In non-relativistic quantum mechanics, the rule that all integer-spin particles are bosons, whereas all half-integer spin particles are fermions, must be accepted as an empirical fact. However, in relativistic quantum mechanics, it is possible to prove that half-integer-spin particles cannot be bosons, and integer-spin particles cannot be fermions. Incidentally, electrons, protons, and neutrons are all fermions.

The *Pauli exclusion principle* is an immediate consequence of the fact that electrons obey Fermi-Dirac statistics. This principle states that no two electrons in a multi-electron system can possess identical sets of observable eigenvalues. For instance, in the case of a three-electron system, the state ket is [see Equation (10.43)]

$$\begin{aligned}|k'\,k''\,k'''\rangle_- = \frac{1}{\sqrt{3!}}\,&(|k'\rangle\,|k''\rangle\,|k'''\rangle + |k'''\rangle\,|k'\rangle\,|k''\rangle + |k''\rangle\,|k'''\rangle\,|k'\rangle\\ &-|k'''\rangle\,|k''\rangle\,|k'\rangle - |k'\rangle\,|k'''\rangle\,|k''\rangle - |k''\rangle\,|k'\rangle\,|k'''\rangle)\,.\end{aligned} \tag{10.44}$$

Note, however, that

$$|k'\,k'\,k'''\rangle = |k'\,k''\,k''\rangle = |k'''\,k''\,k'''\rangle = |0\rangle. \tag{10.45}$$

In other words, if two of the electrons in the system possess the same set of observable eigenvalues then the state ket becomes the null ket, which corresponds to the absence of a state.

10.3 Two-Electron System

Consider a system consisting of two electrons. Let $\mathbf{x}_1$ and $\mathbf{S}_1$ represent the position and spin operators of the first electron, respectively, and let $\mathbf{x}_2$ and $\mathbf{S}_2$ represent the corresponding operators for the second electron. Furthermore, let $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$ represent the total spin operator for the system. Suppose that the Hamiltonian commutes with S^2, as is often the case. It follows that the state of the system is specified by the position eigenvalues $\mathbf{x}_1'$ and $\mathbf{x}_2'$, as well as the total spin quantum numbers s and m. As usual, the eigenvalue of S^2 is $s\,(s+1)\,\hbar^2$, and the eigenvalue of S_z

is $m\hbar$. Moreover, s, m can only takes the values $1, 1$ or $1, 0$ or $1, -1$ or $0, 0$. (See Chapter 6.) The overall wavefunction of the system can be written

$$\psi(\mathbf{x}_1', \mathbf{x}_2'; s, m) = \phi(\mathbf{x}_1', \mathbf{x}_2')\,\chi(s, m), \tag{10.46}$$

where

$$\chi(1, 1) = \chi_+ \chi_+, \tag{10.47}$$

$$\chi(1, 0) = \frac{1}{\sqrt{2}}\,(\chi_+ \chi_- + \chi_- \chi_+), \tag{10.48}$$

$$\chi(1, -1) = \chi_- \chi_-, \tag{10.49}$$

$$\chi(0, 0) = \frac{1}{\sqrt{2}}\,(\chi_+ \chi_- - \chi_- \chi_+). \tag{10.50}$$

Here, the spinor $\chi_+ \chi_-$ denotes a state in which $m_1 = 1/2$ and $m_2 = -1/2$, etc., where $m_1\,\hbar$ and $m_2\,\hbar$ are the eigenvalues of $S_{1\,z}$ and $S_{2\,z}$, respectively. The three $s = 1$ spinors are usually referred to as *triplet* spinors, whereas the single $s = 0$ spinor is called the *singlet* spinor. Note that the triplet spinors are all symmetric with respect to exchange of particles, whereas the singlet spinor is antisymmetric.

Fermi-Dirac statistics requires the overall wavefunction to be antisymmetric with respect to exchange of particles. Now, according to (10.46), the overall wavefunction can be written as a product of a spatial wavefunction and a spinor. Moreover, when the system is in the spin triplet state (i.e., $s = 1$) the spinor is symmetric with respect to exchange of particles. On the other hand, when the system is in the spin singlet state (i.e., $s = 0$) the spinor is antisymmetric. It follows that, to maintain the overall antisymmetry of the wavefunction, the triplet spatial wavefunction must be antisymmetric with respect to exchange of particles, whereas the singlet spatial wavefunction must be symmetric. In other words, in the spin triplet state, the spatial wavefunction takes the form

$$\phi(\mathbf{x}_1', \mathbf{x}_2') = \frac{1}{\sqrt{2}}\left[\omega_A(\mathbf{x}_1')\,\omega_B(\mathbf{x}_2') - \omega_B(\mathbf{x}_1')\,\omega_A(\mathbf{x}_2')\right], \tag{10.51}$$

whereas in the spin singlet state the spatial wavefunction is written

$$\phi(\mathbf{x}_1', \mathbf{x}_2') = \frac{1}{\sqrt{2}}\left[\omega_A(\mathbf{x}_1')\,\omega_B(\mathbf{x}_2') + \omega_B(\mathbf{x}_1')\,\omega_A(\mathbf{x}_2')\right]. \tag{10.52}$$

The probability of observing one electron in the volume element d^3x_1 around position $\mathbf{x}_1$, and the other in the volume element d^3x_2 around position $\mathbf{x}_2$, is $|\phi(\mathbf{x}_1, \mathbf{x}_2)|^2\,d^3x_1\,d^3x_2$, or

$$\frac{1}{2}\left\{|\omega_A(\mathbf{x}_1)|^2\,|\omega_B(\mathbf{x}_2)|^2 + |\omega_A(\mathbf{x}_2)|^2\,|\omega_B(\mathbf{x}_1)|^2 \pm 2\,\mathrm{Re}\left[\omega_A(\mathbf{x}_1)\,\omega_B(\mathbf{x}_2)\,\omega_A^*(\mathbf{x}_2)\,\omega_B^*(\mathbf{x}_1)\right]\right\} d^3x_1\,d^3x_2. \tag{10.53}$$

Here, the plus sign corresponds to the spin singlet state, whereas the minus sign corresponds to the spin triplet state. We can immediately see that in the spin triplet state the probability of finding the two electrons at the same point in space is zero. In other words, the two electrons have a tendency to avoid one another in the triplet state. On the other hand, in the spin singlet state there is an enhanced probability of finding the two electrons at the same point in space, because of the final term in the previous expression. In other words, the two electrons are attracted to one another in the singlet state. Note, however, that the spatial probability distributions associated with the singlet and triplet states only differ substantially when the two single particle spatial wavefunctions $\omega_A(\mathbf{x})$ and $\omega_B(\mathbf{x})$ overlap: i.e., when there exists a region of space in which the two wavefunctions are simultaneously nonnegligible.

10.4 Helium Atom

Consider the helium atom, which is a good example of a two-electron system. The Hamiltonian is written

$$H = \frac{\mathbf{p}_1^2}{2\,m_1} + \frac{\mathbf{p}_2^2}{Z\,m_2} - \frac{Z\,e^2}{4\pi\,\epsilon_0\,r_1} - \frac{Z\,e^2}{4\pi\,\epsilon_0\,r_2} + \frac{e^2}{4\pi\,\epsilon_0\,r_{12}}, \tag{10.54}$$

where $Z = 2$, $r_1 = |\mathbf{x}_1|$, $r_2 = |\mathbf{x}_2|$, and $r_{12} = |\mathbf{x}_1 - \mathbf{x}_2|$. Suppose that the final term on the right-hand side of the above expression were absent. In this case, the overall spatial wavefunction can be formed from products of hydrogen atom wavefunctions calculated with $Z = 2$, instead of $Z = 1$. Each of these wavefunctions is characterized by the usual triplet of quantum numbers, n, l, and m. Now, the total spin of the system is a constant of the motion (since $\mathbf{S}$ obviously commutes with the Hamiltonian), so the overall spin state is either the singlet or the triplet state. The corresponding spatial wavefunction is symmetric in the former case, and antisymmetric in the latter. Suppose that one electron has the quantum numbers n, l, m whereas the other has the quantum numbers n' ,l', m'. The corresponding spatial wavefunction is

$$\phi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}}\left[\psi_{nlm}(\mathbf{x}_1)\,\psi_{n'l'm'}(\mathbf{x}_2) \pm \psi_{nlm}(\mathbf{x}_2)\,\psi_{n'l'm'}(\mathbf{x}_1)\right], \tag{10.55}$$

where the plus and minus signs correspond to the singlet and triplet spin states, respectively. Here, $\psi_{nlm}(\mathbf{x})$ is a standard hydrogen atom wavefunction (calculated with $Z = 2$). For the special case in which the two sets of spatial quantum numbers, n, l, m and n', l', m', are the same, the triplet spin state does not exist (because the associated spatial wavefunction is null). Hence, only singlet spin state is allowed, and the spatial wavefunction reduces to

$$\phi(\mathbf{x}_1, \mathbf{x}_2) = \psi_{nlm}(\mathbf{x}_1)\,\psi_{nlm}(\mathbf{x}_2). \tag{10.56}$$

In particular, the ground state ($n = n' = 1$, $l = l' = 0$, $m = m' = 0$) can only exist as a singlet spin state (i.e., a state of overall spin 0), and has the spatial wavefunction

$$\phi(\mathbf{x}_1, \mathbf{x}_2) = \psi_{100}(\mathbf{x}_1)\,\psi_{100}(\mathbf{x}_2) = \frac{Z^3}{\pi\,a_0^3}\,\exp\left[\frac{-Z\,(r_1 + r_2)}{a_0}\right], \tag{10.57}$$

where a_0 is the Bohr radius. This follows because

$$\psi_{100}(\mathbf{x}) = \frac{1}{\sqrt{4\pi}}\left(\frac{Z}{a_0}\right)^{3/2} 2\,\exp\left(\frac{-Z\,r}{a_0}\right). \tag{10.58}$$

The energy of this state is

$$E = 2\,Z^2\,E_0 = -108.8\,\text{eV}, \tag{10.59}$$

where $E_0 = -13.6\,\text{eV}$ is the ground state energy of a hydrogen atom. In the above expression, the factor of 2 comes from the fact that there are two electrons in a helium atom.

The above estimate for the ground state energy of a helium atom completely ignores the final term on the right-hand side of Equation (10.54), which describes the mutual interaction between the two electrons. We can obtain a better estimate for the ground state energy by treating (10.57) as the unperturbed wavefunction, and $e^2/(4\pi\,\epsilon_0\,r_{12})$ as a perturbation. According to standard first-order perturbation theory, the correction to the ground state energy is

$$\mathit{\Delta}E = \left\langle\frac{e^2}{4\pi\,\epsilon_0\,r_{12}}\right\rangle = \int\!\!\int d^3x_1\,d^3x_2\,\frac{Z^6}{\pi^2\,a_0^6}\,\exp\left[\frac{-2\,Z\,(r_1+r_2)}{a_0}\right]\frac{e^2}{4\pi\,\epsilon_0\,r_{12}}. \tag{10.60}$$

This can be written

$$\frac{\mathit{\Delta}E}{|E_0|} = \frac{2\,Z^6}{\pi^2}\int\!\!\int\frac{d^3x_1}{a_0^3}\,\frac{d^3x_2}{a_0^3}\,\exp\left(-2\,Z\,\frac{r_1+r_2}{a_0}\right)\frac{a_0}{r_{12}}, \tag{10.61}$$

since $E_0 = -e^2/(8\pi\,\epsilon_0\,a_0)$. Now,

$$\frac{1}{r_{12}} = \frac{1}{(r_1^2+r_2^2-2\,r_1\,r_2\,\cos\gamma)^{1/2}} = \sum_{l=0,\infty}\frac{r_<^l}{r_>^{l+1}}\,P_l(\cos\gamma), \tag{10.62}$$

where $r_>$ ($r_<$) is the larger (smaller) of r_1 and r_2, and γ is the angle subtended between $\mathbf{x}_1$ and $\mathbf{x}_2$. Moreover, the so-called *addition theorem* for spherical harmonics states that

$$P_l(\cos\gamma) = \frac{4\pi}{2\,l+1}\sum_{m=-l,+l} Y_{l\,m}^*(\theta_1,\varphi_1)\,Y_{l\,m}(\theta_2,\varphi_2). \tag{10.63}$$

However,

$$\int d\mathit{\Omega}\,Y_{l\,m}(\theta,\phi) = \sqrt{4\pi}\,\delta_{l\,0}\,\delta_{m\,0}, \tag{10.64}$$

so we obtain

$$\begin{aligned}\frac{\mathit{\Delta}E}{|E_0|} &= 32\,Z\int_0^\infty dx_1\,x_1^2\left(\int_0^{x_1} dx_2\,\frac{x_2^2}{x_1}\,\mathrm{e}^{-2\,(x_1+x_2)} + \int_{x_1}^\infty dx_2\,x_2\,\mathrm{e}^{-2\,(x_1+x_2)}\right)\\ &= \frac{5\,Z}{4} = \frac{5}{2}.\end{aligned} \tag{10.65}$$

Here, $x_1 = Z\,r_1/a_0$ and $x_2 = Z\,r_2/a_0$, and $Z = 2$. Thus, our improved estimate for the ground state energy of the helium atom is

$$E = \left(8 - \frac{5}{2}\right)E_0 = -74.8\,\text{eV}. \tag{10.66}$$

This is much closer to the experimental value of $-78.8\,\text{eV}$ than our previous estimate.

Consider an excited state of the helium atom in which one electron is in the ground state, while the other is in a state characterized by the quantum numbers n, l, m. We can write the energy of this state as

$$E = Z^2\,E_{100} + Z^2\,E_{nlm} + \varDelta E, \tag{10.67}$$

where E_{nlm} is the energy of a hydrogen atom electron whose quantum numbers are n, l, m. According to first-order perturbation theory, $\varDelta E$ is the expectation value of $e^2/(4\pi\,\epsilon_0\,r_{12})$. It follows from (10.55) (with $n, l, m = 1, 0, 0$ and $n', l', m' = n, l, m$) that

$$\varDelta E = I \pm J, \tag{10.68}$$

where

$$I = \int d^3x_1 \int d^3x_2\,|\psi_{100}(\mathbf{x}_1)|^2\,|\psi_{nlm}(\mathbf{x}_2)|^2\,\frac{e^2}{4\pi\,\epsilon_0\,r_{12}}, \tag{10.69}$$

$$J = \int d^3x_1 \int d^3x_2\,\psi_{100}(\mathbf{x}_1)\,\psi_{nlm}(\mathbf{x}_2)\,\frac{e^2}{4\pi\,\epsilon_0\,r_{12}}\,\psi^{\,*}_{100}(\mathbf{x}_2)\,\psi^{\,*}_{nlm}(\mathbf{x}_1). \tag{10.70}$$

Here, the plus sign in (10.68) corresponds to the spin singlet state, whereas the minus sign corresponds to the spin triplet state. The integral I—which is known as the *direct integral*—is obviously positive. The integral J—which is known as the *exchange integral*—can be shown to also be positive. Hence, we conclude that in excited states of helium the spin singlet state has a higher energy than the spin triplet state. Incidentally, helium in the spin singlet state is known as *para-helium*, whereas helium in the triplet state is called *ortho-helium*. As we have seen, for the ground state, only para-helium is possible.

The fact that para-helium energy levels lie slightly above corresponding ortho-helium levels is interesting because our original Hamiltonian does not depend on spin. Nevertheless, there is a spin dependent effect—i.e., a helium atom has a lower energy when its electrons possess parallel spins—as a consequence of Fermi-Dirac statistics. To be more exact, the energy is lower in the spin triplet state because the corresponding spatial wavefunction is antisymmetric, causing the electrons to tend to avoid one another (thereby reducing their electrostatic repulsion).

Exercises

10.1 Demonstrate that the particle interchange operator, P_{12}, in a system of two identical particles is Hermitian.

10.2 Consider two identical spin-1/2 particles of mass m confined in a cubic box of dimension L. Find the possible energies and wavefunctions of this system in the case of no interaction between the particles.

10.3 Consider a system of two spin-1 particles with no orbital angular momentum (i.e., both particles are in s-states). What are the possible eigenvalues of the total spin angular momentum of the system, as well as its projection along the z-direction, in the cases in which the particles are non-identical and identical?

11 Relativistic Electron Theory

11.1 Introduction

The aim of this chapter is to develop a quantum mechanical theory of electron dynamics that is consistent with special relativity. Such a theory is needed to explain the origin of electron spin (which is essentially a relativistic effect), and to account for the fact that the spin contribution to the electron's magnetic moment is twice what we would naively expect by analogy with (non-relativistic) classical physics (see Section 5.5). Relativistic electron theory is also required to fully understand the fine structure of the hydrogen atom energy levels (recall, from Section 7.7, and Exercises 7.3 and 7.4, that the modification to the energy levels due to spin-orbit coupling is of the same order of magnitude as the first-order correction due to the electron's relativistic mass increase.)

In the following, we shall use x^1, x^2, x^3 to represent the Cartesian coordinates x, y, z, respectively, and x^0 to represent $c\,t$. The time dependent wavefunction then takes the form $\psi(x^0, x^1, x^2, x^3)$. Adopting standard relativistic notation, we write the four x's as x^μ, for $\mu = 0, 1, 2, 3$. A space-time vector with four components that transforms under Lorentz transformation in an analogous manner to the four space-time coordinates x^μ is termed a *4-vector*, and its components are written like a^μ (i.e., with an upper Greek suffix). We can lower the suffix according to the rules

$$a_0 = a^0, \tag{11.1}$$

$$a_1 = -a^1, \tag{11.2}$$

$$a_2 = -a^2, \tag{11.3}$$

$$a_3 = -a^3. \tag{11.4}$$

Here, the a^μ are called the *contravariant* components of the vector a, whereas the a_μ are termed the *covariant* components. Two 4-vectors a^μ and b^μ have the Lorentz invariant scalar product

$$a^0\,b^0 - a^1\,b^1 - a^2\,a^2 - a^3\,b^3 = a^\mu\,b_\mu = a_\mu\,b^\mu, \tag{11.5}$$

a summation being implied over a repeated letter suffix. The metric tenor $g_{\mu\nu}$ is defined

$$g_{00} = 1, \tag{11.6}$$

$$g_{11} = -1, \tag{11.7}$$

$$g_{22} = -1, \tag{11.8}$$

$$g_{33} = -1, \tag{11.9}$$

with all other components zero. Thus,

$$a_\mu = g_{\mu\nu}\,a^\nu. \tag{11.10}$$

Likewise,

$$a^\mu = g^{\mu\nu}\, a_\nu, \tag{11.11}$$

where $g^{00} = 1$, $g^{11} = g^{22} = g^{33} = -1$, with all other components zero. Finally, $g_\nu^{\;\mu} = g^\mu_{\;\nu} = 1$ if $\mu = \nu$, and $g_\nu^{\;\mu} = g^\mu_{\;\nu} = 0$ otherwise.

In the Schrödinger representation, the momentum of a particle, whose components are written p_x, p_y, p_z, or p^1, p^2, p^3, is represented by the operators

$$p^i = -\mathrm{i}\,\hbar\,\frac{\partial}{\partial x^i}, \tag{11.12}$$

for $i = 1, 2, 3$. Now, the four operators $\partial/\partial x^\mu$ form the covariant components of a 4-vector whose contravariant components are written $\partial/\partial x_\mu$. So, to make expression (11.12) consistent with relativistic theory, we must first write it with its suffixes balanced,

$$p^i = \mathrm{i}\,\hbar\,\frac{\partial}{\partial x_i}, \tag{11.13}$$

and then extend it to the complete 4-vector equation

$$p^\mu = \mathrm{i}\,\hbar\,\frac{\partial}{\partial x_\mu}. \tag{11.14}$$

According to standard relativistic theory, the new operator $p^0 = \mathrm{i}\,\hbar\,\partial/\partial x_0$, which forms a 4-vector when combined with the momenta p^i, is interpreted as the energy of the particle divided by c, where c is the velocity of light in vacuum.

11.2 Dirac Equation

Consider the motion of an electron in the absence of an electromagnetic field. In classical relativity, electron energy, E, is related to electron momentum, $\mathbf{p}$, according to the well-known formula

$$\frac{E}{c} = (p^2 + m_e^2\, c^2)^{1/2}, \tag{11.15}$$

where m_e is the electron rest mass. The quantum mechanical equivalent of this expression is the wave equation

$$\left[p^0 - (p^1\,p^1 + p^2\,p^2 + p^3\,p^3 + m_e^2\,c^2)^{1/2}\right]\psi = 0, \tag{11.16}$$

where the p's are interpreted as differential operators according to Equation (11.14). The above equation takes into account the correct relativistic relation between electron energy and momentum, but is nevertheless unsatisfactory from the point of view of relativistic theory, because it is highly asymmetric between p^0 and the other p's. This makes the equation difficult to generalize, in a manifestly Lorentz invariant manner, in the presence of an electromagnetic field. We must therefore look for a new equation.

If we multiply the wave equation (11.16) by the operator $\left[p^0 + (p^1\,p^1 + p^2\,p^2 + p^3\,p^3 + m_e^2\,c^2)^{1/2}\right]$ then we obtain

$$\left(p^0\,p^0 - p^1\,p^1 - p^2\,p^2 - p^3\,p^3 + m_e^2\,c^2\right)\psi = \left(p^\mu\,p_\mu + m_e^2\,c^2\right)\psi. \tag{11.17}$$

This equation is manifestly Lorentz invariant, and, therefore, forms a more convenient starting point for relativistic quantum mechanics. Note, however, that Equation (11.17) is not entirely equivalent to Equation (11.16), because, although every solution of (11.16) is also a solution of (11.17), the converse is not true. In fact, only those solutions of (11.17) belonging to positive values of p^0 are also solutions of (11.16).

The wave equation (11.17) is quadratic in p^0, and is thus not of the form required by the laws of quantum theory. (Recall that we showed, from general principles, in Chapter 3, that the time evolution equation for the wavefunction should be linear in the operator $\partial/\partial t$, and, hence, in p^0.) We, therefore, seek a wave equation that is equivalent to (11.17), but is linear in p^0. In order to ensure that this equation transforms in a simple way under a Lorentz transformation, we shall require it to be rational and linear in p^1, p^2, p^3, as well as p^0. We are thus lead to a wave equation of the form

$$\left(p^0 - \alpha_1\,p^1 - \alpha_2\,p^2 - \alpha_3\,p^3 - \beta\,m_e\,c\right)\psi = 0, \tag{11.18}$$

where the α's and β are dimensionless, and independent of the p's. Moreover, according to standard relativity, because we are considering the case of no electromagnetic field, all points in space-time must be equivalent. Hence, the α's and β must also be independent of the x's. This implies that the α's and β commute with the p's and the x's. We, therefore, deduce that the α's and β describe an internal degree of freedom that is independent of space-time coordinates. Actually, we shall show later that these operators are related to electron spin.

Multiplying (11.18) by the operator $p^0 + \alpha_1\,p^1 + \alpha_2\,p^2 + \alpha_3\,p^3 + \beta\,m_e\,c$, we obtain

$$\left[p^0\,p^0 - \frac{1}{2}\sum_{i,j=1,3}\{\alpha_i,\alpha_j\}\,p^i\,p^j - \sum_{i=1,3}\{\alpha_i,\beta\}\,p^i\,m_e\,c - \beta^2\,m_e^2\,c^2\right]\psi = 0, \tag{11.19}$$

where $\{a,b\} \equiv a\,b + b\,a$. This equation is equivalent to (11.17) provided that

$$\{\alpha_i,\alpha_j\} = 2\,\delta_{ij}, \tag{11.20}$$

$$\{\alpha_i,\beta\} = 0, \tag{11.21}$$

$$\beta^2 = 1, \tag{11.22}$$

for $i, j = 1, 3$. It is helpful to define the γ^μ, for $\mu = 0, 3$, where

$$\beta = \gamma^0, \tag{11.23}$$

$$\alpha_i = \gamma^0\,\gamma^i, \tag{11.24}$$

for $i = 1, 3$. Equations (11.20)–(11.22) can then be shown to reduce to

$$\{\gamma^\mu,\gamma^\nu\} = 2\,g^{\mu\nu}. \tag{11.25}$$

One way of satisfying the above anti-commutation relations is to represent the operators γ^μ as matrices. However, it turns out that the smallest dimension in which the γ^μ can be realized is four. In fact, it is easily verified that the 4×4 matrices

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{11.26}$$

$$\gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \tag{11.27}$$

for $i = 1, 3$, satisfy the appropriate anti-commutation relations. Here, 0 and 1 denote 2×2 null and identity matrices, respectively, whereas the σ_i represent the 2×2 Pauli matrices introduced in Section 5.7. It follows from (11.23) and (11.24) that

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{11.28}$$

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}. \tag{11.29}$$

Note that γ^0, β, and the α_i, are all Hermitian matrices, whereas the γ^μ, for $\mu = 1, 3$, are anti-Hermitian. However, the matrices $\gamma^0\,\gamma^\mu$, for $\mu = 0, 3$, are Hermitian. Moreover, it is easily demonstrated that

$$\gamma^{\mu\,\dagger} = \gamma^0\,\gamma^\mu\,\gamma^0, \tag{11.30}$$

for $\mu = 0, 3$.

Equation (11.18) can be written in the form

$$(\gamma^\mu\,p_\mu - m_e\,c)\,\psi = (\mathrm{i}\,\hbar\,\gamma^\mu\,\partial_\mu - m_e\,c)\,\psi = 0, \tag{11.31}$$

where $\partial_\mu \equiv \partial/\partial x^\mu$. Alternatively, we can write

$$\mathrm{i}\,\hbar\,\frac{\partial\psi}{\partial t} = (c\,\boldsymbol{\alpha}\cdot\mathbf{p} + \beta\,m_e\,c^2)\,\psi, \tag{11.32}$$

where $\mathbf{p} = (p_x, p_y, p_z) = (p^1, p^2, p^3)$, and $\boldsymbol{\alpha}$ is the vector of the α_i matrices. The previous expression is known as the *Dirac equation*. Incidentally, it is clear that, corresponding to the four rows and columns of the γ^μ matrices, the wavefunction ψ must take the form of a 4×1 column matrix, each element of which is, in general, a function of the x^μ. We saw in Section 5.7 that the spin of the electron requires the wavefunction to have two components. The reason our present theory requires the wavefunction to have four components is because the wave equation (11.17) has twice as many solutions as it ought to have, half of them corresponding to negative energy states.

We can incorporate an electromagnetic field into the above formalism by means of the standard prescription $E \rightarrow E + e\,\phi$, and $p^i \rightarrow p^i + e\,A^i$, where e is the magnitude of the electron charge, ϕ the scalar potential, and $\mathbf{A}$ the vector potential. This prescription can be expressed in the Lorentz invariant form

$$p^\mu \rightarrow p^\mu + \frac{e}{c}\,\Phi^\mu, \tag{11.33}$$

where $\Phi^\mu = (\phi, c\,\mathbf{A})$ is the potential 4-vector. Thus, Equation (11.31) becomes

$$\left[\gamma^\mu \left(p_\mu + \frac{e}{c}\,\Phi_\mu\right) - m_e\,c\right]\psi = \left[\gamma^\mu \left(\mathrm{i}\,\hbar\,\partial_\mu + \frac{e}{c}\,\Phi_\mu\right) - m_e\,c\right]\psi = 0, \tag{11.34}$$

whereas Equation (11.32) generalizes to

$$\mathrm{i}\,\hbar\,\frac{\partial\psi}{\partial t} = \left[-e\,\phi + c\,\boldsymbol{\alpha}\cdot(\mathbf{p} + e\,\mathbf{A}) + \beta\,m_e\,c^2\right]\psi = 0. \tag{11.35}$$

If we write the wavefunction in the spinor form

$$\psi = \begin{pmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} \tag{11.36}$$

then the Hermitian conjugate of Equation (11.35) becomes

$$-\mathrm{i}\,\hbar\,\frac{\partial\psi^\dagger}{\partial t} = \psi^\dagger\left[-e\,\phi + c\,\boldsymbol{\alpha}\cdot(\mathbf{p} + \mathbf{e}\,\mathbf{A}) + \beta\,m_e\,c^2\right] = 0, \tag{11.37}$$

where

$$\psi^\dagger = (\psi_0^*, \psi_1^*, \psi_2^*, \psi_3^*)\,, \tag{11.38}$$

Here, use has been made of the fact that the α_i and β are Hermitian matrices that commute with the p^i and A^i.

It follows from $\psi^\dagger\,\gamma^0$ times Equation (11.34) that

$$\psi^\dagger\left[\gamma^0\,\gamma^\mu\left(\mathrm{i}\,\hbar\,\partial_\mu - \frac{e}{c}\,\Phi_\mu\right) - \gamma^0\,m_e\,c\right]\psi = 0. \tag{11.39}$$

The Hermitian conjugate of this expression is

$$\psi^\dagger\left[\left(-\mathrm{i}\,\hbar\,\partial_\mu - \frac{e}{c}\,\Phi_\mu\right)\gamma^0\,\gamma^\mu - m_e\,c\,\gamma^0\right]\psi = 0, \tag{11.40}$$

where ∂_μ now acts backward on $\psi^\dagger$, and use has been made of the fact that the matrices $\gamma^0\,\gamma^\mu$ and γ^0 are Hermitian. Taking the difference between the previous two equation, we obtain

$$\partial_\mu\,j^\mu = 0, \tag{11.41}$$

where

$$j^\mu = c\,\psi^\dagger\,\gamma^0\,\gamma^\mu\,\psi. \tag{11.42}$$

Writing $j^\mu = (c\,\rho, \mathbf{j})$, where

$$\rho = \psi^\dagger\,\psi, \tag{11.43}$$

$$j^i = c\,\psi^\dagger\,\gamma^0\,\gamma^i\,\psi = \psi^\dagger\,c\,\alpha_i\,\psi, \tag{11.44}$$

Equation (11.41) becomes

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = 0. \tag{11.45}$$

The above expression has the same form as the non-relativistic probability conservation equation (3.65). This suggests that we can interpret the positive definite real scalar field $\rho(\mathbf{x},t) = |\psi|^2$ as the relativistic *probability density*, and the vector field $\mathbf{j}(\mathbf{x},t)$ as the relativistic *probability current*. Integration of the above expression over all space, assuming that $|\psi(\mathbf{x},t)| \rightarrow 0$ as $|\mathbf{x}| \rightarrow \infty$, yields

$$\frac{d}{dt}\int d^3x\;\rho(\mathbf{x},t) = 0. \tag{11.46}$$

This ensures that if the wavefunction is properly normalized at time $t = 0$, such that

$$\int d^3x\;\rho(\mathbf{x},0) = 1, \tag{11.47}$$

then the wavefunction remains properly normalized at all subsequent times, as it evolves in accordance with the Dirac equation. In fact, if this were not the case then it would be impossible to interpret ρ as a probability density. Now, relativistic invariance demands that if the wavefunction is properly normalized in one particular inertial frame then it should be properly normalized in all inertial frames. This is the case provided that Equation (11.41) is Lorentz invariant (i.e., if it has the property that if it holds in one inertial frame then it holds in all inertial frames), which is true as long as the j^{μ} transform as the contravariant components of a 4-vector under Lorentz transformation (see Exercise 11.4).

11.3 Lorentz Invariance of Dirac Equation

Consider two inertial frames, S and S'. Let the x^{μ} and $x^{\mu'}$ be the space-time coordinates of a given event in each frame, respectively. These coordinates are related via a Lorentz transformation, which takes the general form

$$x^{\mu'} = a^{\mu}_{\ \nu}\, x^{\nu}, \tag{11.48}$$

where the $a^{\mu}_{\ \nu}$ are real numerical coefficients that are independent of the x^{μ}. We also have

$$x_{\mu'} = a_{\mu}^{\ \nu}\, x_{\nu}. \tag{11.49}$$

Now, since [see Equation (11.5)]

$$x^{\mu'}\, x_{\mu'} = x^{\mu}\, x_{\mu}, \tag{11.50}$$

it follows that

$$a^{\mu}_{\ \nu}\, a_{\mu}^{\ \lambda} = g_{\nu}^{\ \lambda}. \tag{11.51}$$

Moreover, it is easily shown that

$$x^{\mu} = a_{\nu}^{\ \mu}\, x^{\nu'}, \tag{11.52}$$

$$x_{\mu} = a^{\nu}_{\ \mu}\, x_{\nu'}. \tag{11.53}$$

By definition, a 4-vector p^{μ} has analogous transformation properties to the x^{μ}. Thus,

$$p^{\mu'} = a^{\mu}_{\ \nu}\, p^{\nu}, \tag{11.54}$$

$$p^{\mu} = a_{\nu}^{\ \mu}\, p^{\nu'}, \tag{11.55}$$

etc.

In frame S, the Dirac equation is written

$$\left[\gamma^{\mu}\left(p_{\mu} - \frac{e}{c}\,\Phi_{\mu}\right) - m_e\, c\right]\psi = 0. \tag{11.56}$$

Let ψ' be the wavefunction in frame S'. Suppose that

$$\psi' = A\,\psi, \tag{11.57}$$

where A is a 4×4 transformation matrix that is independent of the x^{μ}. (Hence, A commutes with the p_{μ} and the Φ_{μ}.) Multiplying (11.56) by A, we obtain

$$\left[A\,\gamma^{\mu}\,A^{-1}\left(p_{\mu} - \frac{e}{c}\,\Phi_{\mu}\right) - m_e\, c\right]\psi' = 0. \tag{11.58}$$

Hence, given that the p_{μ} and Φ_{μ} are the covariant components of 4-vectors, we obtain

$$\left[A\,\gamma^{\mu}\,A^{-1}\,a^{\nu}_{\ \mu}\left(p_{\nu'} - \frac{e}{c}\,\Phi_{\nu'}\right) - m_e\, c\right]\psi' = 0. \tag{11.59}$$

Suppose that

$$A\,\gamma^{\mu}\,A^{-1}\,a^{\nu}_{\ \mu} = \gamma^{\nu}, \tag{11.60}$$

which is equivalent to

$$A^{-1}\,\gamma^{\nu}\,A = a^{\nu}_{\ \mu}\,\gamma^{\mu}. \tag{11.61}$$

Here, we have assumed that the $a^{\nu}_{\ \mu}$ commute with A and the γ^{μ} (since they are just numbers). If (11.60) holds then (11.59) becomes

$$\left[\gamma^{\mu}\left(p_{\mu'} - \frac{e}{c}\,\Phi_{\mu'}\right) - m_e\, c\right]\psi' = 0. \tag{11.62}$$

A comparison of this equation with (11.56) reveals that the Dirac equation takes the same form in frames S and S'. In other words, the Dirac equation is Lorentz invariant. Incidentally, it is clear from (11.56) and (11.62) that the γ^{μ} matrices are the same in all inertial frames.

It remains to find a transformation matrix A that satisfies (11.61). Consider an infinitesimal Lorentz transformation, for which

$$a_{\mu}^{\ \nu} = g_{\mu}^{\ \nu} + \Delta\omega_{\mu}^{\ \nu}, \tag{11.63}$$

where the $\Delta\omega_{\mu}^{\ \nu}$ are real numerical coefficients that are independent of the x^{μ}, and are also small compared to unity. To first order in small quantities, (11.51) yields

$$\Delta\omega^{\mu\nu} + \Delta\omega^{\nu\mu} = 0. \tag{11.64}$$

Let us write

$$A = 1 - \frac{\mathrm{i}}{4}\,\sigma_{\mu\nu}\,\Delta\omega^{\mu\nu}, \tag{11.65}$$

where the $\sigma_{\mu\nu}$ are $O(1)$ 4×4 matrices. To first order in small quantities,

$$A^{-1} = 1 + \frac{\mathrm{i}}{4}\,\sigma_{\mu\nu}\,\Delta\omega^{\mu\nu}. \tag{11.66}$$

Moreover, it follows from (11.64) that

$$\sigma_{\mu\nu} = -\sigma_{\nu\mu}. \tag{11.67}$$

To first order in small quantities, Equations (11.61), (11.63), (11.65), and (11.66) yield

$$\Delta\omega^{\nu}{}_{\beta}\,\gamma^{\beta} = -\frac{\mathrm{i}}{4}\,\Delta\omega^{\alpha\beta}\left(\gamma^{\nu}\,\sigma_{\alpha\beta} - \sigma_{\alpha\beta}\,\gamma^{\nu}\right). \tag{11.68}$$

Hence, making use of the symmetry property (11.64), we obtain

$$\Delta\omega^{\alpha\beta}\,(g^{\nu}{}_{\alpha}\,\gamma_{\beta} - g^{\nu}{}_{\beta}\,\gamma_{\alpha}) = -\frac{\mathrm{i}}{2}\,\Delta\omega^{\alpha\beta}\,(\gamma^{\nu}\,\sigma_{\alpha\beta} - \sigma_{\alpha\beta}\,\gamma^{\nu}), \tag{11.69}$$

where $\gamma_{\mu} = g_{\mu\nu}\,\gamma^{\nu}$. Since this equation must hold for arbitrary $\Delta\omega^{\alpha\beta}$, we deduce that

$$2\,\mathrm{i}\,(g^{\nu}{}_{\alpha}\,\gamma_{\beta} - g^{\nu}{}_{\beta}\,\gamma_{\alpha}) = [\gamma^{\nu}, \sigma_{\alpha\beta}]. \tag{11.70}$$

Making use of the anti-commutation relations (11.25), it can be shown that a suitable solution of the above equation is

$$\sigma_{\mu\nu} = \frac{\mathrm{i}}{2}\,[\gamma_{\mu}, \gamma_{\nu}]. \tag{11.71}$$

Hence,

$$A = 1 + \frac{1}{8}\,[\gamma_{\mu}, \gamma_{\nu}]\,\Delta\omega^{\mu\nu}, \tag{11.72}$$

$$A^{-1} = 1 - \frac{1}{8}\,[\gamma_{\mu}, \gamma_{\nu}]\,\Delta\omega^{\mu\nu}. \tag{11.73}$$

Now that we have found the correct transformation rules for an infinitesimal Lorentz transformation, we can easily find those for a finite transformation by building it up from a large number of successive infinitesimal transforms.

Making use of (11.30), as well as $\gamma^{0}\,\gamma^{0} = 1$, the Hermitian conjugate of (11.72) can be shown to take the form

$$A^{\dagger} = 1 - \frac{1}{8}\,\gamma^{0}\,[\gamma_{\mu}, \gamma_{\nu}]\,\gamma^{0}\,\Delta\omega^{\mu\nu} = \gamma^{0}\,A^{-1}\,\gamma^{0}. \tag{11.74}$$

Hence, (11.61) yields

$$A^{\dagger}\,\gamma^{0}\,\gamma^{\mu}\,A = a^{\mu}{}_{\nu}\,\gamma^{0}\,\gamma^{\nu}. \tag{11.75}$$

It follows that

$$\psi^\dagger A^\dagger \gamma^0 \gamma^\mu A \psi = a^\mu_{\ \nu} \psi^\dagger \gamma^0 \gamma^\nu \psi, \tag{11.76}$$

or

$$\psi'^\dagger \gamma^0 \gamma^\mu \psi' = a^\mu_{\ \nu} \psi^\dagger \gamma^0 \gamma^\nu \psi, \tag{11.77}$$

which implies that

$$j^{\mu'} = a^\mu_{\ \nu} j^\nu, \tag{11.78}$$

where the j^μ are defined in Equation (11.42). This proves that the j^μ transform as the contravariant components of a 4-vector.

11.4 Free Electron Motion

According to Equation (11.32), the relativistic Hamiltonian of a free electron takes the form

$$H = c\,\boldsymbol{\alpha}\cdot\mathbf{p} + \beta\, m_e\, c^2. \tag{11.79}$$

Let us use the Heisenberg picture to investigate the motion of such an electron. For the sake of brevity, we shall omit the suffix t that should be appended to dynamical variables that vary in time, according to the formalism of Section 3.2.

The above Hamiltonian is independent of $\mathbf{x}$. Hence, momentum $\mathbf{p}$ commutes with the Hamiltonian, and is therefore a constant of the motion. The x component of the velocity is

$$\dot{x} = \frac{[x, H]}{\mathrm{i}\,\hbar} = c\,\alpha_1, \tag{11.80}$$

where use has been made of the standard commutation relations between position and momentum operators. This result is rather surprising, since it implies a relationship between velocity and momentum that is quite different from that in classical mechanics. This relationship, however, is clearly connected to the expression $j_x = \psi^\dagger c\,\alpha_1\,\psi$ for the x component of the probability current. The operator $\dot{x}$, specified in the above equation, has the eigenvalues $\pm c$, corresponding to the eigenvalues ± 1 of α_1. Since $\dot{y}$ and $\dot{z}$ are similar to $\dot{x}$, we conclude that a measurement of a velocity component of a free electron is certain to yield the result $\pm c$. As is easily demonstrated, this conclusion also holds in the presence of an electromagnetic field.

Of course, electrons are often observed to have velocities considerably less than that of light. Hence, the previous conclusion seems to be in conflict with experimental observations. The conflict is not real, however, because the theoretical velocity discussed above is the velocity at one instance in time, whereas observed velocities are always averages over a finite time interval. We shall find, on further examination of the equations of motion, that the velocity of a free electron is not constant, but oscillates rapidly about a mean value that agrees with the experimentally observed value.

In order to understand why a measurement of a velocity component must lead to the result $\pm c$ in a relativistic theory, consider the following argument. To measure the velocity we must measure the position at two slightly different times, and then divide the change in position by the time interval. (We cannot just measure the momentum and then apply a formula, because the ordinary connection between velocity and momentum is no longer valid.) In order that our measured velocity may approximate to the instantaneous velocity, the time interval between the two measurements of position must be very short, and the measurements themselves very accurate. However, the great accuracy with which the position of the electron is known during the time interval leads to an almost complete indeterminacy in its momentum, according to the Heisenberg uncertainty principle. This means that almost all values of the momentum are equally likely, so that the momentum is almost certain to be infinite. But, an infinite value of a momentum component corresponds to the values $\pm c$ for the corresponding velocity component.

Let us now examine how the election velocity varies in time. We have

$$\mathrm{i}\,\hbar\,\dot{\alpha}_1 = \alpha_1\,H - H\,\alpha_1. \tag{11.81}$$

Now α_1 anti-commutes with all terms in H except $c\,\alpha_1\,p^1$, so

$$\alpha_1\,H + H\,\alpha_1 = \alpha_1\,c\,\alpha_1\,p^1 + c\,\alpha_1\,p^1\,\alpha_1 = 2\,c\,p^1. \tag{11.82}$$

Here, use has been made of the fact that α_1 commutes with p^1, and also that $\alpha_1^2 = 1$. Hence, we get

$$\mathrm{i}\,\hbar\,\dot{\alpha}_1 = 2\,\alpha_1\,H - 2\,c\,p^1. \tag{11.83}$$

Since H and p^1 are constants of the motion, this equation yields

$$\mathrm{i}\,\hbar\,\ddot{\alpha}_1 = 2\,\dot{\alpha}_1\,H, \tag{11.84}$$

which can be integrated to give

$$\dot{\alpha}_1(t) = \dot{\alpha}_1(0)\,\exp\left(\frac{-2\,\mathrm{i}\,H\,t}{\hbar}\right). \tag{11.85}$$

It follows from Equation (11.80) and (11.83) that

$$\dot{x}(t) = c\,\alpha_1(t) = \frac{\mathrm{i}\,\hbar\,c}{2}\,\dot{\alpha}_1(0)\,\exp\left(\frac{-2\,\mathrm{i}\,H\,t}{\hbar}\right)H^{-1} + c^2\,p_x\,H^{-1}, \tag{11.86}$$

and

$$x(t) = x(0) - \frac{\hbar^2\,c}{4}\,\dot{\alpha}_1(0)\,\exp\left(\frac{-2\,\mathrm{i}\,H\,t}{\hbar}\right)H^{-2} + c^2\,p_x\,H^{-1}\,t. \tag{11.87}$$

We can see that the x-component of velocity consists of two parts, a constant part, $c^2\,p_x\,H^{-1}$, connected with the momentum according to the classical relativistic formula, and an oscillatory part whose frequency, $2\,H/h$, is high (being at least $2\,m_e\,c^2/h$). Only the constant part would be observed in a practical measurement of velocity (i.e., an average over a short time interval that is still much longer than $h/2\,m_e\,c^2$). The oscillatory part ensures that the instantaneous value of $\dot{x}$ has the eigenvalues $\pm c$. Note, finally, that the oscillatory part of x is small, being of order $\hbar/m_e\,c$.

11.5 Electron Spin

According to Equation (11.35), the relativistic Hamiltonian of an electron in an electromagnetic field is

$$H = -e\,\phi + c\,\boldsymbol{\alpha}\cdot(\mathbf{p} + e\,\mathbf{A}) + \beta\,m_e\,c^2. \tag{11.88}$$

Hence,

$$\left(\frac{H}{c} + \frac{e}{c}\,\phi\right)^2 = [\boldsymbol{\alpha}\cdot(\mathbf{p} + e\,\mathbf{A}) + \beta\,m_e\,c]^2 = [\boldsymbol{\alpha}\cdot(\mathbf{p} + e\,\mathbf{A})]^2 + m_e^2\,c^2, \tag{11.89}$$

where use has been made of Equations (11.21) and (11.22). Now, we can write

$$\alpha_i = \gamma^5\,\Sigma_i, \tag{11.90}$$

for $i = 1, 3$, where

$$\gamma^5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{11.91}$$

and

$$\Sigma_i = \begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix}. \tag{11.92}$$

Here, 0 and 1 denote 2×2 null and identity matrices, respectively, whereas the σ_i are conventional 2×2 Pauli matrices. Note that $\gamma^5\,\gamma^5 = 1$, and

$$[\gamma^5, \Sigma_i] = 0. \tag{11.93}$$

It follows from (11.89) that

$$\left(\frac{H}{c} + \frac{e}{c}\,\phi\right)^2 = [\boldsymbol{\Sigma}\cdot(\mathbf{p} + e\,\mathbf{A})]^2 + m_e^2\,c^2. \tag{11.94}$$

Now, a straightforward generalization of Equation (5.92) gives

$$(\boldsymbol{\Sigma}\cdot\mathbf{a})\,(\boldsymbol{\Sigma}\cdot\mathbf{b}) = \mathbf{a}\cdot\mathbf{b} + \mathrm{i}\,\boldsymbol{\Sigma}\cdot(\mathbf{a}\times\mathbf{b}), \tag{11.95}$$

where **a** and **b** are any two three-dimensional vectors that commute with $\boldsymbol{\Sigma}$. It follows that

$$[\boldsymbol{\Sigma}\cdot(\mathbf{p} + e\,\mathbf{A})]^2 = (\mathbf{p} + e\,\mathbf{A})^2 + \mathrm{i}\,\boldsymbol{\Sigma}\cdot(\mathbf{p} + e\,\mathbf{A})\times(\mathbf{p} + e\,\mathbf{A}). \tag{11.96}$$

However,

$$(\mathbf{p} + e\,\mathbf{A})\times(\mathbf{p} + e\,\mathbf{A}) = e\,\mathbf{p}\times\mathbf{A} + e\,\mathbf{A}\times\mathbf{p} = -\mathrm{i}\,e\,\hbar\,\nabla\times\mathbf{A} - \mathrm{i}\,e\,\hbar\,\mathbf{A}\times\nabla = -\mathrm{i}\,e\,\hbar\,\mathbf{B}, \tag{11.97}$$

where $\mathbf{B} = \nabla\times\mathbf{A}$ is the magnetic field strength. Hence, we obtain

$$\left(\frac{H}{c} + \frac{e}{c}\,\phi\right)^2 = (\mathbf{p} + e\,\mathbf{A})^2 + m_e^2\,c^2 + e\,\hbar\,\boldsymbol{\Sigma}\cdot\mathbf{B}. \tag{11.98}$$

Consider the non-relativistic limit. In this case, we can write

$$H = m_e\,c^2 + \delta H, \tag{11.99}$$

where δH is small compared to $m_e\,c^2$. Substituting into (11.98), and neglecting δH^2, and other terms involving c^{-2}, we get

$$\delta H \simeq -e\,\phi + \frac{1}{2\,m_e}\,(\mathbf{p} + e\,\mathbf{A})^2 + \frac{e\,\hbar}{2\,m_e}\,\mathbf{\Sigma}\cdot\mathbf{B}. \tag{11.100}$$

This Hamiltonian is the same as the classical Hamiltonian of a non-relativistic electron, except for the final term. This term may be interpreted as arising from the electron having an intrinsic magnetic moment

$$\boldsymbol{\mu} = -\frac{e\,\hbar}{2\,m_e}\,\mathbf{\Sigma}. \tag{11.101}$$

In order to demonstrate that the electron's intrinsic magnetic moment is associated with an intrinsic angular momentum, consider the motion of an electron in a central electrostatic potential: i.e., $\phi = \phi(r)$ and $\mathbf{A} = \mathbf{0}$. In this case, the Hamiltonian (11.88) becomes

$$H = -e\,\phi(r) + c\,\gamma^5\,\mathbf{\Sigma}\cdot\mathbf{p} + \beta\,m_e\,c^2. \tag{11.102}$$

Consider the x component of the electron's orbital angular momentum,

$$L_x = y\,p_z - z\,p_y = \mathrm{i}\,\hbar\left(z\,\frac{\partial}{\partial y} - y\,\frac{\partial}{\partial z}\right). \tag{11.103}$$

The Heisenberg equation of motion for this quantity is

$$\mathrm{i}\,\hbar\,\dot{L}_x = [L_x, H]. \tag{11.104}$$

However, it is easily demonstrated that

$$[L_x, r] = 0, \tag{11.105}$$

$$[L_x, p_x] = 0, \tag{11.106}$$

$$[L_x, p_y] = \mathrm{i}\,\hbar\,p_z, \tag{11.107}$$

$$[L_x, p_z] = -\mathrm{i}\,\hbar\,p_y. \tag{11.108}$$

Hence, we obtain

$$[L_x, H] = \mathrm{i}\,\hbar\,c\,\gamma^5\,(\Sigma_2\,p_z - \Sigma_3\,p_y), \tag{11.109}$$

which implies that

$$\dot{L}_x = c\,\gamma^5\,(\Sigma_2\,p_z - \Sigma_3\,p_y). \tag{11.110}$$

It can be seen that L_x is not a constant of the motion. However, the x-component of the total angular momentum of the system must be a constant of the motion (because a central electrostatic potential exerts zero torque on the system). Hence, we deduce that the electron possesses additional angular momentum that is not connected with its motion through space. Now,

$$\mathrm{i}\,\hbar\,\dot{\Sigma}_1 = [\Sigma_1, H]. \tag{11.111}$$

However,

$$[\Sigma_1, \gamma^5] = 0, \tag{11.112}$$

$$[\Sigma_1, \Sigma_1] = 0, \tag{11.113}$$

$$[\Sigma_1, \Sigma_2] = 2\,\mathrm{i}\,\Sigma_3, \tag{11.114}$$

$$[\Sigma_1, \Sigma_3] = -2\,\mathrm{i}\,\Sigma_2, \tag{11.115}$$

so

$$[\Sigma_1, H] = 2\,\mathrm{i}\,c\,\gamma^5\,(\Sigma_3\,p_y - \Sigma_2\,p_z), \tag{11.116}$$

which implies that

$$\frac{\hbar}{2}\,\dot{\Sigma}_1 = -c\,\gamma^5\,(\Sigma_2\,p_z - \Sigma_3\,p_y). \tag{11.117}$$

Hence, we deduce that

$$\dot{L}_x + \frac{\hbar}{2}\,\dot{\Sigma}_1 = 0. \tag{11.118}$$

Since there is nothing special about the x direction, we conclude that the vector $\mathbf{L} + (\hbar/2)\,\boldsymbol{\Sigma}$ is a constant of the motion. We can interpret this result by saying that the electron has a spin angular momentum $\mathbf{S} = (\hbar/2)\,\boldsymbol{\Sigma}$, which must be added to its orbital angular momentum in order to obtain a constant of the motion. According to (11.101), the relationship between the electron's spin angular momentum and its intrinsic (i.e., non-orbital) magnetic moment is

$$\boldsymbol{\mu} = -\frac{e\,g}{2\,m_e}\,\mathbf{S}, \tag{11.119}$$

where the gyromagnetic ratio g takes the value

$$g = 2. \tag{11.120}$$

As explained in Section 5.5, this is twice the value one would naively predict by analogy with classical physics.

11.6 Motion in Central Field

To further study the motion of an electron in a central field, whose Hamiltonian is

$$H = -e\,\phi(r) + c\,\boldsymbol{\alpha}\cdot\mathbf{p} + \beta\,m_e\,c^2, \tag{11.121}$$

it is convenient to transform to polar coordinates. Let

$$r = (x^2 + y^2 + z^2)^{1/2}, \tag{11.122}$$

and

$$p_r = \mathbf{x}\cdot\mathbf{p}. \tag{11.123}$$

It is easily demonstrated that

$$[r, p_r] = \mathrm{i}\,\hbar, \tag{11.124}$$

which implies that in the Schrödinger representation

$$p_r = -\mathrm{i}\,\hbar\,\frac{\partial}{\partial r}. \tag{11.125}$$

Now, by symmetry, an energy eigenstate in a central field is a simultaneous eigenstate of the total angular momentum

$$\mathbf{J} = \mathbf{L} + \frac{\hbar}{2}\,\boldsymbol{\Sigma}. \tag{11.126}$$

Furthermore, we know from general principles that the eigenvalues of J^2 are $j\,(j+1)\,\hbar^2$, where j is a positive half-integer (since $j = |l + 1/2|$, where l is the standard non-negative integer quantum number associated with orbital angular momentum.)

It follows from Equation (11.95) that

$$(\boldsymbol{\Sigma}\cdot\mathbf{L})\,(\boldsymbol{\Sigma}\cdot\mathbf{L}) = L^2 + \mathrm{i}\,\boldsymbol{\Sigma}\times(\mathbf{L}\times\mathbf{L}). \tag{11.127}$$

However, because $\mathbf{L}$ is an angular momentum, its components satisfy the standard commutation relations

$$\mathbf{L}\times\mathbf{L} = \mathrm{i}\,\hbar\,\mathbf{L}. \tag{11.128}$$

Thus, we obtain

$$(\boldsymbol{\Sigma}\cdot\mathbf{L})\,(\boldsymbol{\Sigma}\cdot\mathbf{L}) = L^2 - \hbar\,\boldsymbol{\Sigma}\cdot\mathbf{L} = J^2 - 2\,\hbar\,\boldsymbol{\Sigma}\cdot\mathbf{L} - \frac{\hbar^2}{4}\,\Sigma^2. \tag{11.129}$$

However, $\Sigma^2 = 3$, so

$$(\boldsymbol{\Sigma}\cdot\mathbf{L} + \hbar)^2 = J^2 + \frac{1}{4}\,\hbar^2. \tag{11.130}$$

Further application of (11.95) yields

$$(\boldsymbol{\Sigma}\cdot\mathbf{L})\,(\boldsymbol{\Sigma}\cdot\mathbf{p}) = \mathbf{L}\cdot\mathbf{p} + \mathrm{i}\,\boldsymbol{\Sigma}\cdot\mathbf{L}\times\mathbf{p} = \mathrm{i}\,\boldsymbol{\Sigma}\cdot\mathbf{L}\times\mathbf{p}, \tag{11.131}$$

$$(\boldsymbol{\Sigma}\cdot\mathbf{p})\,(\boldsymbol{\Sigma}\cdot\mathbf{L}) = \mathbf{p}\cdot\mathbf{L} + \mathrm{i}\,\boldsymbol{\Sigma}\cdot\mathbf{p}\times\mathbf{L} = \mathrm{i}\,\boldsymbol{\Sigma}\cdot\mathbf{p}\times\mathbf{L}, \tag{11.132}$$

However, it is easily demonstrated from the fundamental commutation relations between position and momentum operators that

$$\mathbf{L}\times\mathbf{p} + \mathbf{p}\times\mathbf{L} = 2\,\mathrm{i}\,\hbar\,\mathbf{p}. \tag{11.133}$$

Thus,

$$(\boldsymbol{\Sigma}\cdot\mathbf{L})\,(\boldsymbol{\Sigma}\cdot\mathbf{p}) + (\boldsymbol{\Sigma}\cdot\mathbf{p})\,(\boldsymbol{\Sigma}\cdot\mathbf{L}) = -2\,\hbar\,\boldsymbol{\Sigma}\cdot\mathbf{p}, \tag{11.134}$$

which implies that

$$\{\boldsymbol{\Sigma}\cdot\mathbf{L} + \hbar,\ \boldsymbol{\Sigma}\cdot\mathbf{p}\} = 0. \tag{11.135}$$

Now, $\gamma^5\,\boldsymbol{\Sigma} = \boldsymbol{\alpha}$. Moreover, γ^5 commutes with $\mathbf{p}$, $\mathbf{L}$, and $\boldsymbol{\Sigma}$. Hence, we conclude that

$$\{\boldsymbol{\Sigma}\cdot\mathbf{L} + \hbar,\ \boldsymbol{\alpha}\cdot\mathbf{p}\} = 0. \tag{11.136}$$

Finally, since β commutes with $\mathbf{p}$ and $\mathbf{L}$, but anti-commutes with the components of $\boldsymbol{\alpha}$, we obtain

$$[\zeta, \boldsymbol{\alpha}\cdot\mathbf{p}] = 0, \tag{11.137}$$

where

$$\zeta = \beta\,(\boldsymbol{\Sigma}\cdot\mathbf{L} + \hbar)\,. \tag{11.138}$$

If we repeat the above analysis, starting at Equation (11.131), but substituting $\mathbf{x}$ for $\mathbf{p}$, and making use of the easily demonstrated result

$$\mathbf{L}\times\mathbf{x} + \mathbf{x}\times\mathbf{L} = 2\,\mathrm{i}\,\hbar\,\mathbf{x}, \tag{11.139}$$

we find that

$$[\zeta, \boldsymbol{\alpha}\cdot\mathbf{x}] = 0. \tag{11.140}$$

Now, r commutes with β, as well as the components of $\boldsymbol{\Sigma}$ and $\mathbf{L}$. Hence,

$$[\zeta, r] = 0. \tag{11.141}$$

Moreover, β commutes with the components of $\mathbf{L}$, and can easily be shown to commute with all components of $\boldsymbol{\Sigma}$. It follows that

$$[\zeta, \beta] = 0. \tag{11.142}$$

Hence, Equations (11.121), (11.137), (11.141), and (11.142) imply that

$$[\zeta, H] = 0. \tag{11.143}$$

In other words, an eigenstate of the Hamiltonian is a simultaneous eigenstate of ζ. Now,

$$\zeta^2 = [\beta\,(\boldsymbol{\Sigma}\cdot\mathbf{L} + \hbar)]^2 = (\boldsymbol{\Sigma}\cdot\mathbf{L} + \hbar)^2 = J^2 + \frac{1}{4}\,\hbar^2, \tag{11.144}$$

where use has been made of Equation (11.130), as well as $\beta^2 = 1$. It follows that the eigenvalues of ζ^2 are $j\,(j+1)\,\hbar^2 + (1/4)\,\hbar^2 = (j+1/2)^2\,\hbar^2$. Thus, the eigenvalues of ζ can be written $k\,\hbar$, where $k = \pm(j + 1/2)$ is a non-zero integer.

Equation (11.95) implies that

$$\begin{aligned}(\boldsymbol{\Sigma}\cdot\mathbf{x})\,(\boldsymbol{\Sigma}\cdot\mathbf{p}) &= \mathbf{x}\cdot\mathbf{p} + \mathrm{i}\,\boldsymbol{\Sigma}\cdot\mathbf{x}\times\mathbf{p} = r\,p_r + \mathrm{i}\,\boldsymbol{\Sigma}\cdot\mathbf{L}\\ &= r\,p_r + \mathrm{i}\,(\beta\,\zeta - \hbar),\end{aligned} \tag{11.145}$$

where use has been made of (11.123) and (11.138).

It is helpful to define the new operator ϵ, where

$$r\,\epsilon = \boldsymbol{\alpha}\cdot\mathbf{x}. \tag{11.146}$$

Moreover, it is evident that

$$[\epsilon, r] = 0. \tag{11.147}$$

Hence,

$$r^2\,\epsilon^2 = (\boldsymbol{\alpha}\cdot\mathbf{x})^2 = \frac{1}{2}\sum_{i,j=1,3}\{\alpha_i, \alpha_j\}\,x^i\,x^j = \sum_{i=1,3}(x^i)^2 = r^2, \tag{11.148}$$

where use has been made of (11.20). It follows that

$$\epsilon^2 = 1. \tag{11.149}$$

We have already seen that ζ commutes with $\boldsymbol{\alpha}\cdot\mathbf{x}$ and r. Thus,

$$[\zeta, \epsilon] = 0. \tag{11.150}$$

Equation (11.95) gives

$$(\boldsymbol{\Sigma}\cdot\mathbf{x})\,(\mathbf{x}\cdot\mathbf{p}) - (\mathbf{x}\cdot\mathbf{p})\,(\boldsymbol{\Sigma}\cdot\mathbf{x}) = \boldsymbol{\Sigma}\cdot\left[\mathbf{x}\,(\mathbf{x}\cdot\mathbf{p}) - (\mathbf{x}\cdot\mathbf{p})\,\mathbf{x}\right] = \mathrm{i}\,\hbar\,\boldsymbol{\Sigma}\cdot\mathbf{x}, \tag{11.151}$$

where use has been made of the fundamental commutation relations for position and momentum operators. However, $\mathbf{x}\cdot\mathbf{p} = r\,p_r$ and $\boldsymbol{\Sigma}\cdot\mathbf{x} = \gamma^5\,r\,\epsilon$, so, multiplying through by γ^5, we get

$$r^2\,\epsilon\,p_r - r\,p_r\,r\,\epsilon = \mathrm{i}\,\hbar\,r\,\epsilon. \tag{11.152}$$

Equation (11.124) then yields

$$[\epsilon, p_r] = 0. \tag{11.153}$$

Equation (11.145) implies that

$$(\boldsymbol{\alpha}\cdot\mathbf{x})\,(\boldsymbol{\alpha}\cdot\mathbf{p}) = r\,p_r + \mathrm{i}\,(\beta\,\zeta - \hbar). \tag{11.154}$$

Making use of Equations (11.141), (11.146), (11.147), and (11.149), we get

$$\boldsymbol{\alpha}\cdot\mathbf{p} = \epsilon\,(p_r - \mathrm{i}\,\hbar/r) + \mathrm{i}\,\epsilon\beta\,\zeta/r. \tag{11.155}$$

Hence, the Hamiltonian (11.121) becomes

$$H = -e\,\phi(r) + c\,\epsilon\,(p_r - \mathrm{i}\,\hbar/r) + \mathrm{i}\,c\,\epsilon\beta\,\zeta/r + \beta\,m_e\,c^2. \tag{11.156}$$

Now, we wish to solve the energy eigenvalue problem

$$H\,\psi = E\,\psi, \tag{11.157}$$

where E is the energy eigenvalue. However, we have already shown that an eigenstate of the Hamiltonian is a simultaneous eigenstate of the ζ operator belonging to the eigenvalue $k\,\hbar$, where k is a non-zero integer. Hence, the eigenvalue problem reduces to

$$\left[-e\,\phi(r) + c\,\epsilon\,(p_r - \mathrm{i}\,\hbar/r) + \mathrm{i}\,c\,\hbar\,k\,\epsilon\beta/r + \beta\,m_e\,c^2\right]\psi = E\,\psi, \tag{11.158}$$

which only involves the radial coordinate r. It is easily demonstrated that ϵ anti-commutes with β. Hence, given that β takes the form (11.28), and that $\epsilon^2 = 1$, we can represent ϵ as the matrix

$$\epsilon = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}. \tag{11.159}$$

Thus, writing ψ in the spinor form

$$\psi = \begin{pmatrix} \psi_a(r) \\ \psi_b(r) \end{pmatrix}, \tag{11.160}$$

and making use of (11.125), the energy eigenvalue problem for an electron in a central field reduces to the following two coupled radial differential equations:

$$\hbar\, c \left(\frac{d}{dr} + \frac{k+1}{r}\right) \psi_b + (E - m_e\, c^2 + e\, \phi)\, \psi_a = 0, \tag{11.161}$$

$$\hbar\, c \left(\frac{d}{dr} - \frac{k-1}{r}\right) \psi_a - (E + m_e\, c^2 + e\, \phi)\, \psi_b = 0. \tag{11.162}$$

11.7 Fine Structure of Hydrogen Energy Levels

For the case of a hydrogen atom,

$$\phi(r) = \frac{e}{4\pi\, \epsilon_0\, r}. \tag{11.163}$$

Hence, Equations (11.161) and (11.162) yield

$$\left(\frac{1}{a_1} - \frac{\alpha}{y}\right) \psi_a - \left(\frac{d}{dy} + \frac{k+1}{y}\right) \psi_b = 0, \tag{11.164}$$

$$\left(\frac{1}{a_2} + \frac{\alpha}{y}\right) \psi_b - \left(\frac{d}{dy} - \frac{k-1}{y}\right) \psi_a = 0, \tag{11.165}$$

where $y = r/a_0$, and

$$a_1 = \frac{\alpha}{1 - \mathcal{E}}, \tag{11.166}$$

$$a_2 = \frac{\alpha}{1 + \mathcal{E}}, \tag{11.167}$$

with $\mathcal{E} = E/(m_e\, c^2)$. Here, $a_0 = 4\pi\, \epsilon_0\, \hbar^2/(m_e\, e^2)$ is the Bohr radius, and $\alpha = e^2/(4\pi\, \epsilon_0\, \hbar\, c)$ the fine structure constant. Writing

$$\psi_a(y) = \frac{\mathrm{e}^{-y/a}}{y}\, f(y), \tag{11.168}$$

$$\psi_b(y) = \frac{\mathrm{e}^{-y/a}}{y}\, g(y), \tag{11.169}$$

where

$$a = (a_1\, a_2)^{1/2} = \frac{\alpha}{\sqrt{1-\mathcal{E}^2}}, \tag{11.170}$$

we obtain

$$\left(\frac{1}{a_1} - \frac{\alpha}{y}\right) f - \left(\frac{d}{dy} - \frac{1}{a} + \frac{k}{y}\right) g = 0, \tag{11.171}$$

$$\left(\frac{1}{a_2} + \frac{\alpha}{y}\right) g - \left(\frac{d}{dy} - \frac{1}{a} - \frac{k}{y}\right) f = 0. \tag{11.172}$$

Let us search for power law solutions of the form

$$f(y) = \sum_s c_s\, y^s, \tag{11.173}$$

$$g(y) = \sum_s c_s'\, y^s, \tag{11.174}$$

where successive values of s differ by unity. Substitution of these solutions into Equations (11.171) and (11.172) leads to the recursion relations

$$\frac{c_{s-1}}{a_1} - \alpha\, c_s - (s+k)\, c_s' + \frac{c_{s-1}}{a} = 0, \tag{11.175}$$

$$\frac{c_{s-1}'}{a_2} + \alpha\, c_s' - (s-k)\, c_s + \frac{c_{s-1}}{a} = 0. \tag{11.176}$$

Multiplying the first of these equations by a, and the second by a_2, and then subtracting, we eliminate both c_{s-1} and c_{s-1}', since $a/a_1 = a_2/a$. We are left with

$$[a\,\alpha - a_2\,(s-k)]\, c_s + [a_2\,\alpha + a\,(s+k)]\, c_s' = 0. \tag{11.177}$$

The physical boundary conditions at $y = 0$ require that $y\,\psi_a \to 0$ and $y\,\psi_b \to 0$ as $y \to 0$. Thus, it follows from (11.168) and (11.169) that $f \to 0$ and $g \to 0$ as $y \to 0$. Consequently, the series (11.173) and (11.174) must terminate at small positive s. If s_0 is the minimum value of s for which c_s and c_s' do not both vanish then it follows from (11.175) and (11.176), putting $s = s_0$ and $c_{s_0-1} = c_{s_0-1}' = 0$, that

$$\alpha\, c_{s_0} + (s_0 + k)\, c_{s_0}' = 0, \tag{11.178}$$

$$\alpha\, c_{s_0}' - (s_0 - k)\, c_{s_0} = 0, \tag{11.179}$$

which implies that

$$\alpha^2 = -s_0^2 + k^2. \tag{11.180}$$

Since the boundary condition requires that the minimum value of s_0 be greater than zero, we must take

$$s_0 = (k^2 - \alpha^2)^{1/2}. \tag{11.181}$$

To investigate the convergence of the series (11.173) and (11.174) at large y, we shall determine the ratio c_s/c_{s-1} for large s. In the limit of large s, Equations (11.176) and (11.177) yield

$$s\,c_s \simeq \frac{c_{s-1}}{a} + \frac{c'_{s-1}}{a_2}, \tag{11.182}$$

$$a_2\,c_s \simeq a\,c'_s, \tag{11.183}$$

since $\alpha \simeq 1/137 \ll 1$. Thus,

$$\frac{c_s}{c_{s-1}} \simeq \frac{2}{a\,s}. \tag{11.184}$$

However, this is the ratio of coefficients in the series expansion of $\exp(2\,y/a)$. Hence, we deduce that the series (11.173) and (11.174) diverge unphysically at large y unless they terminate at large s.

Suppose that the series (11.173) and (11.174) terminate with the terms c_s and c'_s, so that $c_{s+1} = c'_{s+1} = 0$. It follows from (11.175) and (11.176), with $s+1$ substituted for s, that

$$\frac{c_s}{a_1} + \frac{c'_s}{a} = 0, \tag{11.185}$$

$$\frac{c'_s}{a_2} + \frac{c_s}{a} = 0. \tag{11.186}$$

These two expressions are equivalent, because $a^2 = a_1\,a_2$. When combined with (11.177) they give

$$a_1\,[a\,\alpha - a_2\,(s-k)] = a\,[a_2\,\alpha + a\,(s+k)]\,, \tag{11.187}$$

which reduces to

$$2\,a_1\,a_2\,s = a\,(a_1 - a_2)\,\alpha, \tag{11.188}$$

or

$$\mathcal{E} = \left(1 + \frac{\alpha^2}{s^2}\right)^{-1/2}. \tag{11.189}$$

Here, s, which specifies the last term in the series, must be greater than s_0 by some non-negative integer i. Thus,

$$s = i + (k^2 - \alpha^2)^{1/2} = i + [(j+1/2)^2 - \alpha^2]^{1/2}. \tag{11.190}$$

where $j\,(j+1)\,\hbar^2$ is the eigenvalue of J^2. Hence, the energy eigenvalues of the hydrogen atom become

$$\frac{E}{m_e\,c^2} = \left\{1 + \frac{\alpha^2}{(i + [(j+1/2)^2 - \alpha^2]^{1/2})^2}\right\}^{-1/2}. \tag{11.191}$$

Given that $\alpha \simeq 1/137$, we can expand the above expression in α^2 to give

$$\frac{E}{m_e\,c^2} = 1 - \frac{\alpha^2}{2\,n^2} - \frac{\alpha^4}{2\,n^4}\left(\frac{n}{j+1/2} - \frac{3}{4}\right) + O(\alpha^6), \tag{11.192}$$

where $n = i + j + 1/2$ is a positive integer. Of course, the first term in the above expression corresponds to the electron's rest mass energy. The second term corresponds to the standard non-relativistic expression for the hydrogen energy levels, with n playing the role of the radial quantum number (see Section 4.6). Finally, the third term corresponds to the fine structure correction to these energy levels (see Exercise 7.4). Note that this correction only depends on the quantum numbers n and j. Now, we showed in Section 7.7 that the fine structure correction to the energy levels of the hydrogen atom is a combined effect of spin-orbit coupling and the electron's relativistic mass increase. Hence, it is evident that both of these effects are automatically taken into account in the Dirac equation.

11.8 Positron Theory

We have already mentioned that the Dirac equation admits twice as many solutions as it ought to, half of them belonging to states with negative values for the kinetic energy $c\,p^0 + e\,\Phi^0$. This difficulty was introduced when we passed from Equation (11.16) to Equation (11.17), and is inherent in any relativistic theory.

Let us examine the negative energy solutions of the equation

$$\left[\left(p^0 + \frac{e}{c}\,\Phi^0\right) - \alpha_1\left(p^1 + \frac{e}{c}\,\Phi^1\right) - \alpha_2\left(p^2 + \frac{e}{c}\,\Phi^2\right) - \alpha_3\left(p^3 + \frac{e}{c}\,\Phi^3\right) - \beta\,m_e\,c\right]\psi = 0 \qquad (11.193)$$

a little more closely. For this purpose, it is convenient to use a representation of the α's and β in which all the elements of the matrices α_1, α_2, and α_3 are real, and all of those of the matrix representing β are imaginary or zero. Such a representation can be obtained from the standard representation by interchanging the expressions for α_2 and β. If Equation (11.193) is expressed as a matrix equation in this representation, and we then substitute $-\mathrm{i}$ for i, we get [remembering the factor i in Equation (11.14)]

$$\left[\left(p^0 - \frac{e}{c}\,\Phi^0\right) - \alpha_1\left(p^1 - \frac{e}{c}\,\Phi^1\right) - \alpha_2\left(p^2 - \frac{e}{c}\,\Phi^2\right) - \alpha_3\left(p^3 - \frac{e}{c}\,\Phi^3\right) - \beta\,m_e\,c\right]\psi^* = 0. \qquad (11.194)$$

Thus, each solution, ψ, of the wave equation (11.193) has for its complex conjugate, ψ^*, a solution of the wave equation (11.194). Furthermore, if the solution, ψ, of (11.193) belongs to a negative value for $c\,p^0 + e\,\Phi^0$ then the corresponding solution, ψ^*, of (11.194) will belong to a positive value for $c\,p^0 - e\,\Phi^0$. But, the operator in (11.194) is just what we would get if we substituted $-e$ for e in the operator in (11.193). It follows that each negative energy solution of (11.193) is the complex conjugate of a positive energy solution of the wave equation obtained from (11.193) by the substitution of $-e$ for e. The latter solution represents an electron of charge $+e$ (instead of $-e$, as we have had up to now) moving through the given electromagnetic field.

We conclude that the negative energy solutions of (11.193) refer to the motion of a new type of particle having the mass of an electron, but the opposite charge. Such particles have been observed experimentally, and are called *positrons*. Note that we cannot simply assert that the negative energy solutions represent positrons, since this would make the dynamical relations all wrong. For instance, it is certainly not true that a positron has a negative kinetic energy. Instead, we assume that nearly all of the negative energy states are occupied, with one electron in each state, in accordance with the Pauli exclusion principle. An unoccupied negative energy state will now appear as a particle with a positive energy, since to make it disappear we would have to add an electron with a negative energy to the system. We assume that these unoccupied negative energy states correspond to positrons.

The previous assumptions require there to be a distribution of electrons of infinite density everywhere in space. A perfect vacuum is a region of space in which all states of positive energy are unoccupied, and all of those of negative energy are occupied. In such a vacuum, the Maxwell equation

$$\nabla\cdot\mathbf{E} = 0 \tag{11.195}$$

must be valid. This implies that the infinite distribution of negative energy electrons does not contribute to the electric field. Thus, only departures from the vacuum distribution contribute to the electric charge density ρ in the Maxwell equation

$$\nabla\cdot\mathbf{E} = \frac{\rho}{\epsilon_0}. \tag{11.196}$$

In other words, there is a contribution $-e$ for each occupied state of positive energy, and a contribution $+e$ for each unoccupied state of negative energy.

The exclusion principle ordinarily prevents a positive energy electron from making transitions to states of negative energy. However, it is still possible for such an electron to drop into an unoccupied state of negative energy. In this case, we would observe an electron and a positron simultaneously disappearing, their energy being emitted in the form of radiation. The converse process would consist in the creation of an electron positron pair from electromagnetic radiation.

Exercises

11.1 Noting that $\alpha_i = -\beta\,\alpha_i\,\beta$, prove that the α_i and β matrices all have zero trace. Hence, deduce that each of these matrices has n eigenvalues $+1$, and n eigenvalues -1, where $2\,n$ is the dimension of the matrices.

11.2 Verify that the matrices (11.28) and (11.29) satisfy Equations (11.20)–(11.22).

11.3 Verify that the matrices (11.26) and (11.27) satisfy the anti-commutation relations (11.25).

11.4 Verify that if

$$\partial_\mu\, j^\mu = 0,$$

where j^μ is a 4-vector field, then

$$\int d^3x\, j^0$$

is Lorentz invariant, where the integral is over all space, and it is assumed that $j^\mu \to 0$ as $|\mathbf{x}| \to \infty$.

11.5 Verify that (11.71) is a solution of (11.70).

11.6 Verify that the 4×4 matrices Σ_i, defined in (11.92), satisfy the standard anti-commutation relations for Pauli matrices: i.e.,

$$\{\Sigma_i, \Sigma_j\} = 2\,\delta_{ij}.$$